(2)

REPORT DO **AD-A229 032**

DTIC ELECTE NOV 16 1990 S B D

1a. REPORT SECURITY CLASSIFICATION: Unclassified

1b. RESTRICTIVE MARKINGS: N/A

2a. SECURITY CLASSIFICATION AUTHORITY: N/A

3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release Distribution Unlimited

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A

4. PERFORMING ORGANIZATION REPORT NUMBER(S):

5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 90 1100

6a. NAME OF PERFORMING ORGANIZATION: Utah State University

6b. OFFICE SYMBOL (If applicable): N/A

7a. NAME OF MONITORING ORGANIZATION: AFOSR/NL

6c. ADDRESS (City, State and ZIP Code): Logan, UT 84322-8200

7b. ADDRESS (City, State and ZIP Code): Bolling AFB, DC 20332-6448 Bld 410

8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR

8b. OFFICE SYMBOL (If applicable): NL

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: Grant AFOSR - 89-0509

8c. ADDRESS (City, State and ZIP Code): Bld 410 Bolling AFB, DC 20332

10. SOURCE OF FUNDING NOS

| PROGRAM ELEMENT NO | PROJECT NO | TASK NO | WORK UNIT NO |
|---|---|---|---|
| 61102F | 2312 | A4 | |

11. TITLE (Include Security Classification): Environmental Containment Property Estimation Using QSARs in An Expert System

12. PERSONAL AUTHOR(S): Doucette, W.J.; Stevens, D.K.; Dupont, R.R.; McLean, J.E.; Denne, D.; Holt, M.

13a. TYPE OF REPORT: Annual

13b. TIME COVERED: FROM _____ TO _____

14. DATE OF REPORT (Yr., Mo., Day): 1990/9/14

15. PAGE COUNT:

16. SUPPLEMENTARY NOTATION:

| 17 COSATI CODES | | |
|---|---|---|
| FIELD | GROUP | SUB. GR |
| | | |

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Pollution, QSARs, Expert System, Organic Chemical, Property Estimation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A microcomputer based Property Estimation Program (PEP) and Database (DB), utilizing molecular connectivity indices (MCI)-property and property-property correlations, as well as UNIFAC derived activity coefficients, has been designed to provide both experts and non-experts with a fast, economical method to estimate compound aqueous solubility, octanol/water partition coefficient, vapor pressure, organic carbon normalized soil sorption coefficient, BCF, and Henry's Law constant for use in environmental fate modeling. The user can input the required structural information using either Simplified Molecular Input Line Entry System (SMILES) notation or connection tables generated from two commercially available two-dimensional drawing programs, ChemDraw or ChemIntosh . Estimates of predictor accuracy are provided along with the estimated property values. The development and current status of the PEP-DB program is described.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION: Unclass

22a. NAME OF RESPONSIBLE INDIVIDUAL: T. Jan Cerveny, Lt Col, USAF

22b. TELEPHONE NUMBER (Include Area Code): (202) 767-5021

22c. OFFICE SYMBOL: NL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE



17 SEP 1990

# ENVIRONMENTAL CONTAINMENT PROPERTY ESTIMATION USING QSARs IN AN EXPERT SYSTEM

William J. Doucette
David K. Stevens
R. Ryan Dupont
Joan E. McLean
Doug Denne
Mark Holt
Utah State University
Utah Water Research Laboratory
Logan, UT 84322-8200

September 14, 1990

Annual Report for Period 15 August 1989 - 15 August 1990

Prepared for
U.S. AIR FORCE OF SCIENTIFIC RESEARCH
Bolling AFB, OC 20332-6448

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT"D)

## LIST OF TABLES



## LIST OF FIGURES

## EXECUTIVE SUMMARY

In order to assess the potential impact of the accidental introduction of an organic chemical into the environment, information is needed concerning its environmental fate. The fate of an organic chemical in the environment depends on a variety of physical, chemical and biological processes. Mathematical models, which attempt to integrate these processes, are widely used to predict the transport and distribution of organic contaminants in the environment. Use of these models requires a variety of input parameters which describe site and contaminant physical-chemical and biological characteristics. Several important contaminant properties used to assess the mobility and persistence of a chemical are aqueous solubility, octanol/water partition coefficient, soil/water sorption coefficient, Henry's Law constant, bioconcentration factor, and transformation rates for biodegradation, photolysis and hydrolysis.

One major limitation to the use of environmental fate models has been the lack of suitable values for many of these properties. The scarcity of data, due mainly to the difficulty and cost involved in experimental determination of such properties, has resulted in an increased reliance on the use of estimated values for many applications.

Quantitative Structure-Property Relationships (QSPRs) and Quantitative Property-Property Relationships (QPPRs) are methods by which properties of a chemical can be estimated from a knowledge of the structure of a molecule or from another more easily obtained property. Selection and application of the most appropriate QSPRs or QPPRs for a given compound is based on several factors including: the availability of required input, the methodology for calculating the necessary topological information, the appropriateness of a correlation to the chemical of interest, and an understanding of the mechanisms controlling the property being estimated.

A microcomputer based **P**roperty **E**stimation **P**rogram and **D**atabase (PEP-DB), utilizing molecular connectivity indices (MCI)-property and property-property correlations and UNIFAC derived activity coefficients, is being developed to provide both experts and non-

experts with a fast, economical method to estimate a compound's aqueous solubility, octanol/water partition coefficient, vapor pressure, organic carbon normalized soil sorption coefficient ($K_{oc}$), bioconcentration factor (BCF), and Henry's Law constant for use in environmental fate modeling. The user can input the required structural information using either Simplified Molecular Input Line Entry System (SMILES) notation or connection tables generated from two commercially available two-dimensional drawing programs, ChemDraw™ or ChemIntosh™. Estimates of predictor accuracy are provided along with the estimated property values. This report describes the development and current status of PEP-DB.

## OBJECTIVES OR STATEMENT OF WORK-WJD

The primary goal of this project is to develop a microcomputer-based expert system utilizing Quantitative Structure Activity Relationships (QSARs) to predict the physical-chemical properties of an organic chemical which are necessary to model its environmental fate. The specific properties that are being investigated include: aqueous solubility (S), vapor pressure (Vp), organic carbon normalized soil/water partition coefficient ($K_{oc}$), Henry's Law constant (H), and bioconcentration factor (BCF).

In order to achieve the primary goal of this research, the following specific objectives are being accomplished:

1. To compile an accurate database of experimentally determined values of aqueous solubility, vapor pressure, soil/water partition coefficient, Henry's Law constant, and bioconcentration and bioaccumulation factors for a wide variety of organic compounds. The database includes compounds exhibiting a broad range of physical and chemical properties and expected mobility and persistence.
2. Using the database developed in Objective 1, evaluate and refine existing methods and/or develop new methods for estimating these contaminant properties using QSARs.
3 Develop a microcomputer-based decision support system which incorporates the methods developed in Objective 2, to allow the prediction of environmental fate and transport properties of an organic contaminant upon inputting its structure. An estimate of the accuracy of the predicted value is also provided from the decision support system.
4. Test the ability of the decision support system developed in Objective 3 to provide an accurate estimate of these environmental fate and transport properties. This will be done using a test set of chemicals of interest to the USAF (solvents, fuels, pesticides) for which accurate experimental values are available.

## BACKGROUND AND SIGNIFICANCE

### Nature of the problem

In order to assess the potential impact that the introduction of an organic chemical into the environment will have, information is needed concerning its environmental fate. Environmental fate encompasses the transport and degradation processes which determine the behavior of a chemical released into the environment. The fate of an organic chemical introduced into the environment depends on a variety of physical, chemical and biological processes. Mathematical models, which attempt to integrate these processes, are widely used to predict the environmental transport and distribution of organic contaminants. Use of these models requires a variety of input parameters concerning site and contaminant physical-chemical and biological characteristics. Several important contaminant properties used to assess the mobility and persistence of a chemical are listed below:

| Mobility | Persistence |
|---|---|
| Henry's Law constant | Biodegradation Rate |
| (or vapor pressure and aqueous solubility) | Photolysis Rate |
| Bioconcentration factor | Hydrolysis Rate |
| Soil/water partition coefficients | Oxidation Rate |

One major limitation to the use of such models has been the lack of suitable values for many of the properties listed above. The scarcity of data, due mainly to the difficulty and cost involved in experimental determination of such properties, has resulted in an increased reliance on the use of estimated values for many applications.

### Quantitative Structure-activity relationships (QSARs

Quantitative Structure-Activity Relationships (QSARs) are sources of such data that are increasingly recognized as rapid, practical, and inexpensive methods with which to estimate values of some constants or properties necessary for fate assessment models. QSARs are methods by which data or information on the properties of a chemical can be inferred or

calculated from a knowledge of the structure of a molecule or from another more easily obtained property without a specific concern for molecular structure.

Most QSAR methods currently used to estimate contaminant properties fall into one the following categories (Lyman, 1985):

1. Correlations between the property of interest and another more easily obtained property.
2. Correlations between the property of interest and various topological indexes.
3. Calculation of the property of interest using fragment or group contribution methods.
4. Theoretical equations, generally containing parameters that are experimentally or empirically derived.

Correlations between the property of interest and another more easily obtained property:

One of the most useful and widely used type of estimation method is a simple linear regression between two properties. Frequently this regression is expressed in terms of the log of the two properties. Researchers have found that a number of environmental properties can be related to one another in this manner. For example, octanol/water partition coefficient ($K_{ow}$) has been used to estimate soil sorption coefficients (Karickhoff, 1979) aqueous solubility (Chiou et al, 1977 and Mackay et al., 1980), bioconcentration factors (Neely et al., 1987, Chiou et al., 1977), and aquatic toxicity (Koneman, 1980).

One important limitation in using this approach is that in many cases, values for the property used to estimate the property of interest are also not available. In addition, when using this approach it is essential to evaluate the data used to generate the correlation expression. In many instances the reliability correlation expression was derived using only one chemical class, a narrow range of property values, poor quality data, or estimated property values in the regression analysis. It is also important not to use a regression equation outside of the range of data from which it was derived.

Fragment constant methods

These methods generally assume that a single numerical value, referred to as a fragment constant, will represent the contribution of a specified atom, fragment (a group of atoms

bonded together), or structural factor to the property of interest. Probably the most widely used fragment constant method has been developed by Hansch and Leo (1979) for estimating octanol-water partition coefficients ($K_{ow}$). Using a large database of measured values of $K_{ow}$, fragment constants have been developed for over 160 atoms or fragments and for a variety of structural factors (double and triple bond, ring aromatic rings, etc.). These fragment constants and structural factors are used to estimate a value of log $K_{ow}$ for a particular chemical using the following expression.

$$\sum \log K_{ow} = \sum (\text{fragment values}) + \sum (\text{factor values}) \quad (1)$$

Another example of the fragment constant approach to predicting properties is the UNAFAC (UNIQUAC Functional Group Activity Coefficient) solution of groups method of calculating activity coefficients. The UNIFAC method was developed to estimate activity coefficients in mixtures of nonelectrolytes (Fredenslund et al., 1977). In this technique, the activity coefficient is divided into two parts, a combinatorial part which reflects the size and shape of the molecule present and a residual portion which depends on functional group interactions. Various parameters, such as van der Waals group volumes and surface areas and group interaction parameters, are input into a series of equations from which the combinatorial and residual parts are calculated. Values for the group parameters have been tabulated and can be found in the literature (Frendenslund et al., 1977 and Gmehling, 1982). Lyman et al. (1982) give several examples illustrating the use of this technique.

The UNIFAC method was used by Arbunkle (1983) to calculate the activity coefficients for 21 organic compounds. Solubility values were than calculated from the UNIFAC derived activity coefficients and compared to experimental values. The calculated solubility values were generally lower than the experimental values and the largest errors were generally associated with the least soluble compounds.

Lyman (1985) summarized the limitations associated with the use of fragment constant methods as: most fragment constants are derived from compounds with no more that one

functional group, most are limited in the number of fragments and factors they cover, meaning that for certain types of compounds the fragments and factors are unavailable; the application to structurally complex molecules can be difficult; many make no distinction between positional isomers of a compound; and they may be difficult to use.

Correlations Between the Property of Interest and Topological Indexes

Topological indexes attempt to translate molecular structures into unique characteristic structural descriptors that can be expressed numerically. One of the most widely used topological index is the molecular connectivity index (MCI) developed by Randic (1972), and refined and expanded by Kier and Hall (1976, 1980, 1986). Molecular connectivity is a method of bond counting from which topological indexes, based on the structure of the compound, can be derived. For a given molecular structure, several types and orders of molecular connectivity indexes (MCIs) can be calculated. Information on the molecular size, branching, cyclization, unsaturation, and heteroatom content of a molecule is encoded in these various indices (Kier and Hall, 1976). Molecular connectivity has been utilized to predict $K_{oc}$ (Sabljic, 1984; Sabljic, 1987b), S (Kenaga and Goring, 1980), and other physicochemical properties of chemical compounds such as $K_{ow}$ (Doucette and Andren, 1988) and bioconcentration factors (Briggs, 1981). The advantage of using MCIs to predict physical-chemical properties is that once the correlation has been developed only the structure of the chemical of interest is required as input. No additional experimental parameters are needed.

Although researches have been successful in using MCIs to estimate properties for a variety of chemicals, the problem of class specific correlations still remains. For example, Gerstl and Helling (1987) evaluated the use of MCIs in estimating log $K_{oc}$, log $K_{ow}$ and water solubility for many types of pesticides and non-pesticides. It was found that, while good predictions of sorption coefficients were possible for a specific groups of compounds, the ability of any one equation to predict log $K_{oc}$, based upon one or two MCIs, was rather low for diverse compound types. In addition, calculation of MCIs can be difficult, especially the higher-order indices for complex molecules.

Theoretically Derived Equations

Two examples of using theoretical equations to estimate properties of environmental interest include the estimation of a compound's vapor pressure from its boiling point and the calculation of Henry's Law constant from the ratio of a compound's vapor pressure to its aqueous solubility.

Using a 72 compound test set of hydrocarbons and halocarbons, Mackay et al. (1982) developed an expression which enables the estimation of a compound's vapor pressure from its boiling point. This equation was derived, in part, from the Clausius-Clapeyron equation, itself derived from the second law of thermodynamics. Mackay et al. noted that this expression may not be applicable to other classes of compounds and that method errors increase as vapor pressure decreases.

Another widely used example of a theoretically derived method is the calculation of Herny's Law constants, H, from a compound's vapor pressure/aqueous solubility ratio. H is defined as the ratio of a chemicals concentration in air to its concentration in water when those two phases are in contact and at equilibrium. The derivation of this expression requires the assumption that liquid phase activity coefficients are constant up to the aqueous solubility limit. Thus, the method is not applicable to compounds with high water solubilities.

Problems Associated with the Estimation Methods (or the need for a decision support system)

In most cases, more than one estimation method is available for a particular input parameter. Estimation methods however, have widely varying accuracies and indiscriminate use of these techniques can result in large errors.

Selection and application of QSARs methods requires varying degrees of expertise that depend on the structure of a particular chemical of interest, knowledge of the mechanism of the process, the extent of the database used to develop the QSAR, and the complexity of the structural analysis required to relate structure to the property. For example, some QSARs are broader than others in the range of chemicals that are covered, and some methods have been established with a better understanding of the mechanisms or properties involved. In many

cases estimation methods are developed from empirical or semiempirical correlations. The success of the correlation is dependent on many factors including the type and number of compounds used in its development.

Incorporation of QSARs into a computer format is a logical and necessary step to gain full advantage of the methodologies for simplifying fate assessment. A practical computerized property estimation program, utilizing QSARs, should include the following attributes: be simple and flexible to use for both experts and non-experts, include sufficient statistical information regarding the development of the QASARs so that the range of applicability of such models can be evaluated, and provide an indication of the accuracy of the estimated property. A microcomputer-based system for the estimation of parameters necessary for fate assessment models would be of great benefit to USAF agencies responsible for environmental fate assessment.

## STATUS OF RESEARCH EFFORT

### Introduction

After evaluating a variety of software approaches, including several expert systems shells, we decided to build the property estimation system using Apple HyperCard™ software. This approach will enable us to efficiently build a flexible, simple-to-use interface between the various modules or subroutines of our property estimation system. This approach will also permit additional property estimation routines, as they are developed, to be easily added.

The following sections will describe the development and use of the HyperCard™-based Property Estimation Program, PEP, and its associated chemical property Database (DB).

### Overview of PEP and Chemical Property Database

PEP is currently comprised of three property estimation modules linked to a chemical property database. The three property estimation modules utilize MCI-property relationships, UNIFAC-derived activity coefficients, and property-property correlations for compound property estimates. The modular organization of PEP is illustrated in Figure 1.

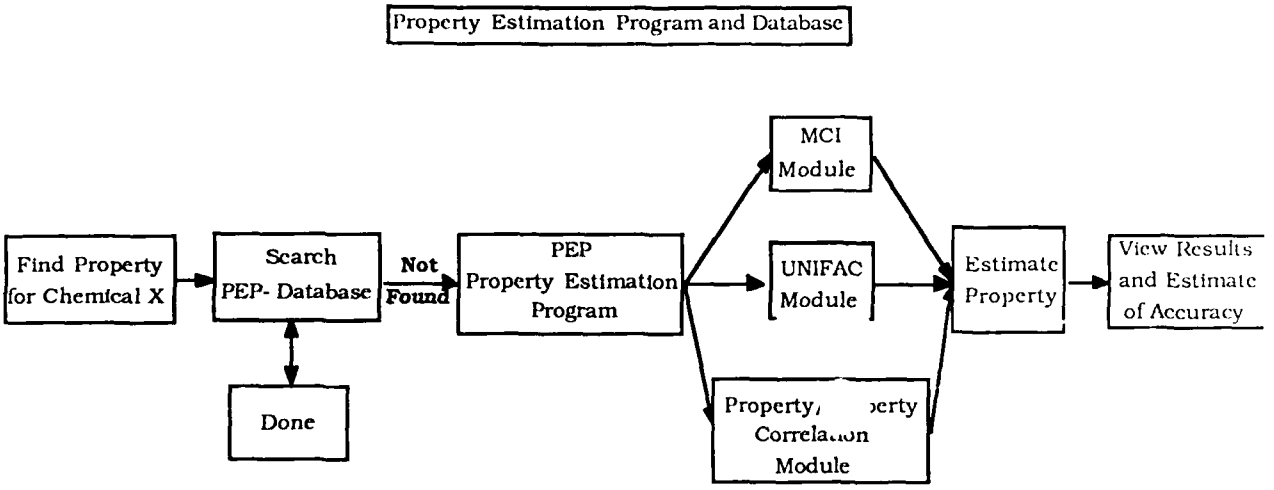



Figure 1. Flow chart overview of PEP.DB

## COMPUTER HARDWARE/SOFTWARE REQUIREMENTS AND DEVELOPMENT TOOLS

### PEP Software Overview

The PEP software system is a HyperCard™ based program that runs on Apple Macintosh computers. HyperCard is an information/management program included with the purchase of Macintosh computers. HyperCard offers graphics, information storage, the means to display information in a variety of formats, the ability to establish links between related information, a high level language (HyperTalk), the ability to extend HyperTalk by writing new commands in a compiled language, and a mechanism to transfer control to other Macintosh applications. The PEP system uses all these features.

### PEP Software Components

PEP uses a variety of programs to control the user interface, manage the HyperCard Stacks, and make various computations. These tools are: HyperTalk Scripts, HyperTalk External Commands (XCMDs), HyperTalk External Functions (XFCNs) and external application programs.

HyperTalk Scripts

HyperCard™ contains a high level, interpreted language called HyperTalk. PEP makes extensive use of HyperTalk. A HyperTalk program is called a script. PEP uses scripts to control the user interface, the stack to stack linkage, and the linkages between the cards in each stack. This is the standard way in which HyperTalk is normally used. Like most HyperCard™ applications, PEP uses scripts at all levels in the HyperCard™ hierarchy. Scripts are used at the button, field, card, background, and stack levels in each of the PEP stacks.

Besides controlling each stack, scripts also do some of the computations in the system. For example, a script is used to compute estimates of the chemical/physical properties in the PEP processors as a function of the MCIs.

HyperCard External Commands and Functions

PEP contains several external commands and functions. A HyperCard™ external command is an extension to the HyperTalk language. HyperCard™ externals in PEP are written in the language C. External commands and functions are used for several reasons: to do functions not supported by HyperTalk, to improve the speed of some computations, and to improve the structure of a software module.

External Applications

External applications are external computer programs that can be run independent of HyperCard™. HyperCard™ provides a means (the open command) to transfer control to another application. When the application ends, control returns to HyperCard™. PEP uses three applications:

1. ChemDraw™, by Cambridge Scientific Computing, Cambridge, Mass.
2. Chemintosh™, by SoftShell International Ltd, Grand Junction, CO
3. EstimateMCI, by Utah Water Research Laboratory, Logan, Utah

ChemDraw™ and Chemintosh™ are commercial applications. They are used by the PEP processor stack to provide a means for the user to create a connection table for subsequent input to PEP. These programs are not distributed with PEP. EstimateMCI is an application

designed and coded by the PEP development team. This is a C program. It is distributed with PEP. Estimate MCI accepts a SMILES string or the contents of a connection table file for its input. It then computes the MCIs as a function of the SMILES string or the connection table. It communicates with the PEP processor stack by passing and receiving information through external files. (See the prologue in the source code file MCI.C for a detailed description of this interface.)

PEP Software Tools

The PEP software uses a five software tools. These are:

1. HyperCard, version 1.2.5, by Apple Computer, Inc.
2. Think C, version 4.0, by Symantec Corporation, Cupertino, California
3. XTRA, the XFCN, XCMD toolkit, by Adrian Freed, Fidcor USA, Louisville, Colorado
4. ResEdit, version 1.2, by Apple Computer, Inc.
5. Progress XCMD, by Jay Hodgdon, 587 Cutwater Lane, Foster City, Ca. 94404

HyperCard™

This is the foundation of the system. All modules in the PEP system are based on HyperCard™.

Think C

All external commands and functions (XCMDs, and XFCNS) are written in the language C, and compiled with version 4.0 of Symantec Corporation's Think C compiler. They are compiled as Macintosh code resources. The application, estimate MCI, was also written in C and compiled with the Think C compiler. This was compiled as a Macintosh application.

XTRA

XTRA is a commercial product that eases the burden of writing HyperTalk external commands and functions. This product provides an interface between HyperCard™ and code resources (external commands or functions) compiled with the Think C compiler. In Macintosh terminology, this interface is called glue. The XTRA program also contains a library of useful functions for use by HyperCard™ external commands.

ResEdit

ResEdit is a Macintosh resource editor by Apple Computer, Inc. ResEdit is used to attach each HyperCard™ external command and function to the appropriate HyperCard stack.

Progress XCMD

The progress XCMD is a shareware product. This XCMD displays a dialog box with a moving cursor to show how far along (e.g. % complete) time consuming scripts while computations are done.

System Requirements

The PEP system requires the following system configuration to run: a Macintosh Plus, Macintosh SE, or Macintosh II computer, with a hard disk; HyperCard software; Macintosh system software version 5.0 or greater, running under MultiFinder; and a minimum of 2 megabytes of memory (RAM), with 1000 kBytes of memory allocated for HyperCard.

## CHEMICAL PROPERTY DATABASE

Description

Experimentally determined chemical property data was complied from a variety of literature sources and computerized databases. Using this information, a chemical property database was developed using HyperCard™. This database was used for developing MCI-property and property-property relationships and is a major component of the overall property estimation software system being developed. In its current state, the database includes the following information: compound name and synonyms, CAS number, chemical formula, molecular weight, boiling point, melting point, aqueous solubility, octanol/water partition coefficient, vapor pressure, soil/water sorption coefficients, Henry's Law constants, bioconcentration factors and appropriate references for each value. The number of compounds currently in the database for each property is summarized in Table 1.

TABLE 1. Current number of compounds in chemical property database listed by property

| Property name | # compounds |
|---|---|
| Aqueous solubility | 365 |
| Octanol/water partition coefficient | 196 |
| Soil sorption coefficient (organic carbon normalized) | 171 |
| Vapor Pressure | 95 |
| Henry's Law constants | 76 |
| Bioconcentration Factors | 70 |

The Chemical Property Database also provides the means for the user to search for chemical compounds, to sort the compounds by name, boiling point, melting point, or molecular weight, and the ability to transfer to any of the PEP modules. The chemical property database screen is illustrated in Figure 2.

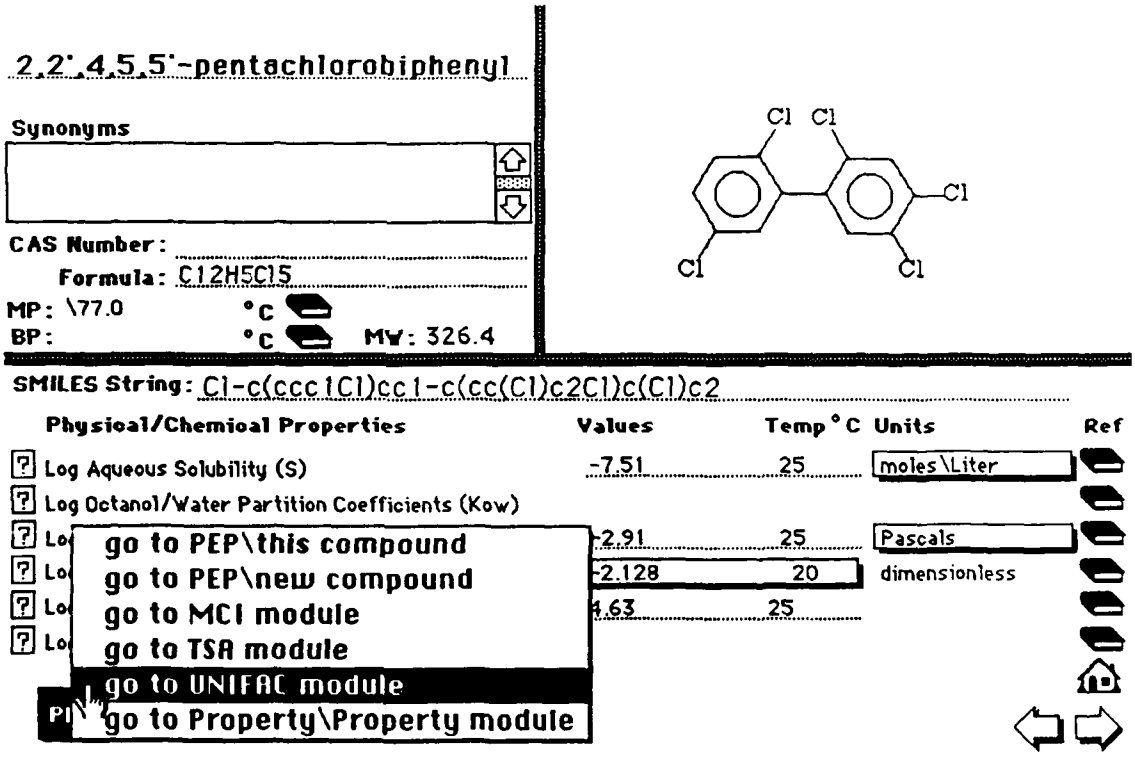



Figure 2. View of PEP's chemical property database

## MCI-BASED PROPERTY ESTIMATION MODULE

### Overview

Upon entering the MCI module the user must first input the necessary structural information. The user can input the required structural information using either SMILES (Simplified Molecular Input Line Entry System) notation or connection tables generated from two commercially available two-dimensional drawing programs, ChemDraw™ or ChemIntosh™. A detailed description of SMILES will be incorporated as a help option in the near future. After the structural information is entered, MCIs are then calculated using an application external to HyperCard™. The calculation of MCIs will be described in detail in a following section. After the MCIs are calculated, the results are imported back into HyperCard™ where they can be displayed. Upon importing the MCI the user can then choose which property are to be estimated. Several MCI-property regression models are available for each property. A view statistics option is available to aid the user in choosing the most appropriate model. After choosing the most appropriate regression, estimates for the selected properties can be made. The MCI module results window provides an estimate of the property along with its calculated accuracy based on the 95% confidence interval calculated from the regression. The overall operation of the MCI module is illustrated in Figure 3.

### Calculation of MCIs

To calculate the MCIs for a given compound, a delta (d) value must first be assigned to each atom in the structure. Three main d values were computed in this study: normal, bond, and valence. Normal deltas were computed by summing the number of bonds (single, double, etc. are counted as one bond) connected to the atom whose delta is being calculated. The bond deltas were calculated the same way as the normal deltas except the bonds were taken at their face value (single is one, double is two, etc.) instead of each bond being equal to one. Valence deltas for each atom were computed according to equations (2) and (3) (Kier and Hall, 1986):

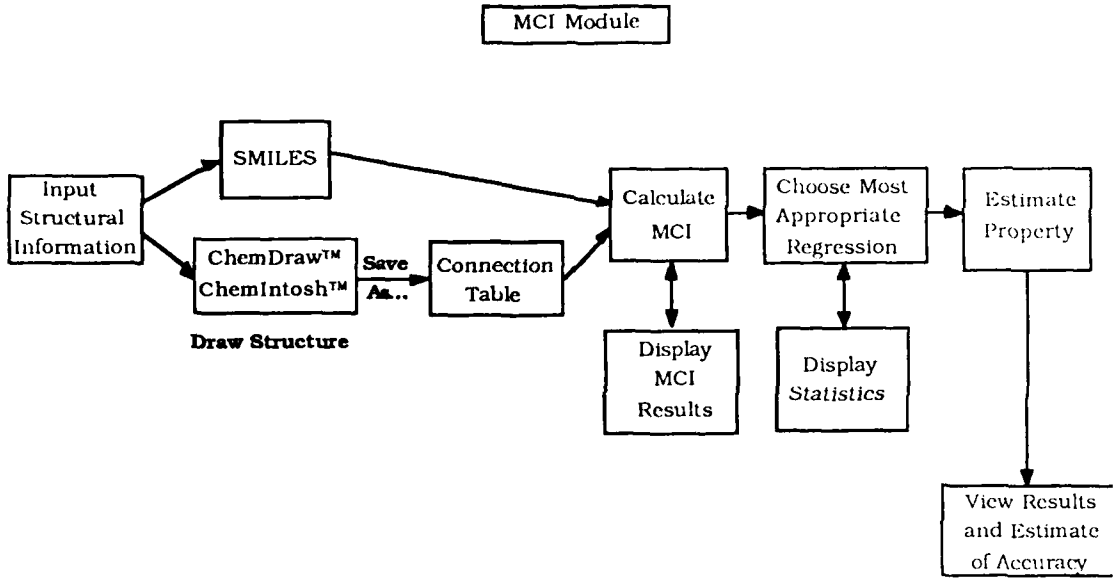



Figure 3. Flow chart depicting operation of MCI module

$$d^v = Z^v - h \tag{2}$$

$$d^v = \backslash f(Z^v - h, Z - Z^v) \tag{3}$$

where $d^v$ is the valence delta, $Z^v$ is the number of valence electrons in the atom, h is the number of hydrogen atoms bound to the atom, and Z is the atomic number of the atom. Equation (1) is used for those atoms in the first row of the periodic chart, and equation (2) is used for all other atoms. An example delta calculation for phenol is shown in Figure 4.

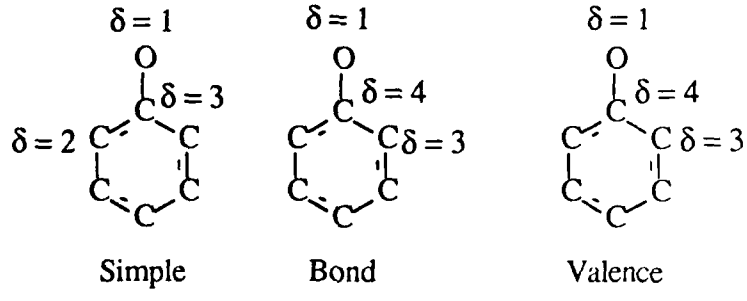



Figure 4. Delta values calculated for phenol.

Once the delta values have been calculated for each atom in the molecule, simple, bond and valence indices of different orders and types can be calculated for a given molecule. The order refers to the number of bonds in the skeletal substructure or fragment used in computing the index: zero order defines individual atoms, first order uses individual bond lengths, second order uses two adjacent bond combinations, and so on. The type refers to the structural fragment (path, cluster, path/cluster or chain) used in computing the index the MCIs corresponding to the desired graph fragment types can be calculated. The fragment types are derived from graph theory and are best described by example, as shown in Figure 5.

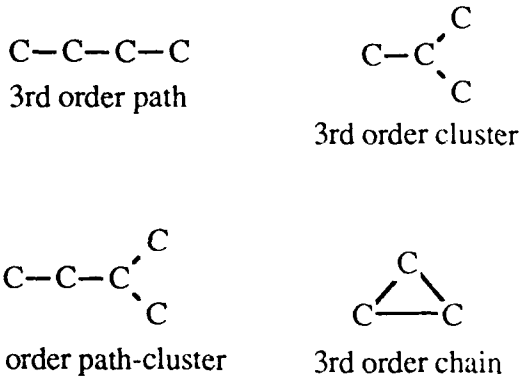



Figure 5. Four types of graph fragments.

Only path indices are possible for orders less than 3. The symbol $^{2}\chi$ represents a simple second order index whereas the symbol $^{1}\chi^{v}$ represents a first order valence index.

Finally, to calculate the MCIs, the following equation is used (Kier, 1980):

$$ {}^{m}\chi_{g}^{t} = \sum_{i=1}^{n} \frac{1}{\sqrt{\prod_{j=1}^{m+1} \delta_{j}^{t}}} \qquad (4) $$

where $\delta^{t}$ is the delta of type t determined as above, n is the total possible number of $m^{th}$ order indices in the molecule, m is the number of bonds over which the deltas are taken, t is the type of indices (normal, bond, or valence), and g is one of the four graph types (path, cluster, path-cluster, or chain).

The MCI calculation routine PEP calculates simple, bond and valence indices of several types (path, cluster, chain, and path/cluster) and orders (0 through 6), if possible, for each molecule, resulting in a maximum of 54 index values for each molecule.

To account for non-dispersive force effects on aqueous solubility and solubility related properties zero through six order Δ valence path indices ($\Delta\chi$), as described by Bahnick and Doucette (1988), are calculated by PEP, in addition to the 54 indices described above. To calculate $\Delta\chi$ indices, a nonpolar equivalent is made by substituting C for O or N atoms. MCIs are calculated for the nonpolar equivalent and values for $\Delta\chi$ can be computed for each type of index by:

$$\Delta\chi = (\chi)_{np} - \chi \qquad (5)$$

## DEVELOPMENT OF MCI-PROPERTY RELATIONSHIPS

For each property, MCI-property relationships were developed for both general and specific chemical classes. The database compounds were classified into four general groups: Non-polar aromatics (compounds having no O or N containing functional groups), polar aromatic (compounds having O or N containing functional groups), non-polar non-cyclic non-aromatics, polar non-cyclic non-aromatics and specific chemical classes such as PCBs, PAHs, carbamates, ureas etc. In addition, "universal" equations were developed which utilized all database compounds having values for a specific property.

Two approaches were used to choose the most appropriate variable(s) in developing the MCI-property regression equations. The first approach used a combination of two indices, one related to molecular size or dispersive intermolecular forces (i.e., vp0, vp1, np0, np1, bp0 and bp1) and one related to the non-dispersive forces ($\Delta\chi$). The second approach relied entirely on a stepwise multiple linear regression program to select the most appropriate variable. If the the two approaches resulted in models of similar fit the equation resulting from the first approach was used because of its greater conceptual meaning.

The MCI-property relationships which have been developed so far are presented in Appendix A along with the relevant regression statistics. The information obtained in the first year of the project shows that MCIs can be used to predict property values for a variety of organic chemical types using both class specific and more general regression equations. The universal and more general regression models utilize $\Delta\chi$ indices. These non-dispensive force terms are important in predicting physical properties for molecules exhibiting substantial hydrophilicity. The predicted property values are within the experimental uncertainty in their measurement for the vast majority of chemicals investigated.

In addition, an improvement in predicted values for some of the compounds in the study could be realized by adjusting the assigned valance values. This will be further investigated in the second year of study.

## STATISTICAL EVALUATION OF MCI-PROPERTY RELATIONSHIPS

All regression equations used in MCI-property relationships (and in the property/property correlations discussed in a later section) were evaluated for their statistical significance of regression variables. Each of these steps were independently performed for each regression relationship that was developed from experimental data.

### Examination of Residuals

Residuals, $e_i$, are defined mathematically as the difference between the $\log_{10}$ of an experimentally determined value, $Y_i$, and its predicted value, $Y_i'$:

$$e_i = \log_{10}(Y_i - Y_i'), \; i = 1, \ldots n_{obs} \tag{6}$$

Regression analyses are performed subject to a number of important assumptions relative to the nature of their residuals (Draper and Smith 1981): 1) they are assumed to be independent of one another in a data set, 2) they are assumed to have a mean of zero, 3) they are assumed to have constant variance, and 4) they are assumed to be normally distributed. If the regression

relationship is a true representation of the experimental data and includes all significant variables, then its residuals should confirm the assumptions made above.

A test of the validity of a developed regression expression can then be based on an assessment of the validity of these assumptions regarding its residuals through an examination of residual plots (residuals versus measured values, residuals versus predicted values, a normal probability plot of residual values) and calculated statistics.

Plots of residual values versus measured and predicted values will generally take the form of the relationships shown in Figure 6. Figure 6a indicates residuals that verify the assumptions made above being independent with constant variance, and normally distributed about a zero mean. The relationship shown in Figure 6b indicates either a lack of independence of the residual values, or a regression model that does not adequately represent observed data, i.e., a linear model not adequately representing a relationship that has curvature. Figure 6c indicates non-constant variance, while the residual plot in Figure 6d indicates error in the analysis such as the absence of the y-intercept term in a regression model. The assumption of normally distributed residuals can also be evaluated from a normal probability plot of residual values. These residuals can be said to be normally distributed if their normal probability plot is linear.

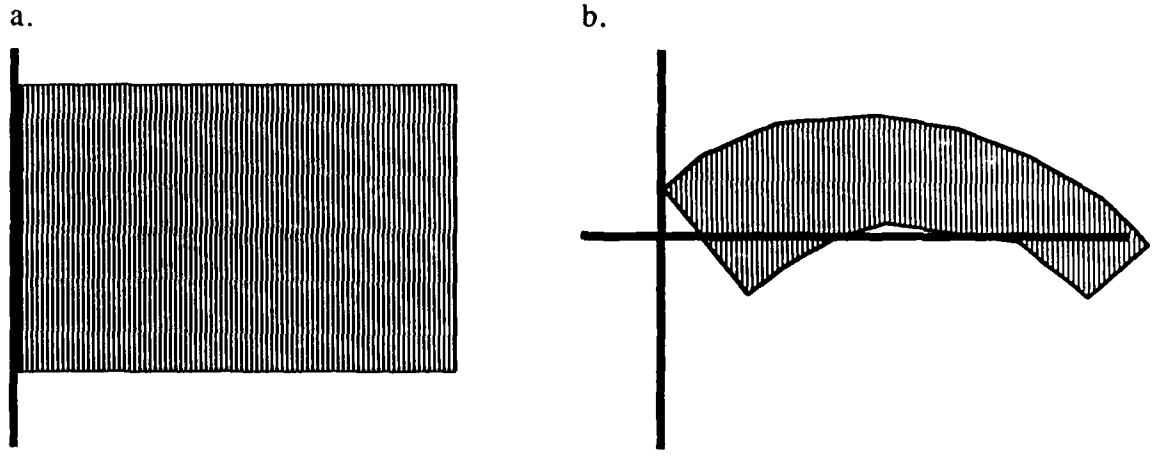



Figure 6. Examples of residual plots.

c. d.

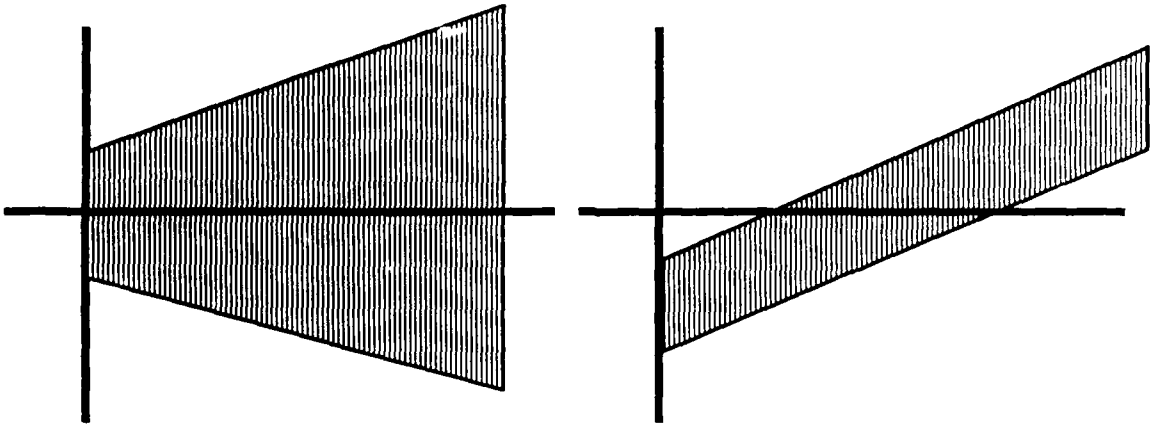


Figure 6. (Cont'd)

Deviations from the "correct" plot behavior, i.e., Figure 6a, does not automatically mean that there is an error in the regression relationship or in the assumptions of normality, constant variance, etc., of the residuals. This is particularly true for regression relationships developed from a small number of experimental observations. Gross deviations from ideal behavior should be identified, and should be flags for further testing of the validity of assumptions regarding the data set being analyzed as described by Anscombe and Tukey (1963) and Draper and Smith (1981).

Analysis of variance

When applied to regression analyses, the analysis of variance (ANOVA) is a test for the significance of the regression relationship, i.e., are the regression variables a significantly better descriptor of the behavior of the data than its mean value. An example ANOVA table is presented in Table 2. In this table, the significance of the regression relationship is indicated by the F ratio, which represents the ratio of the variance explained by the regression to that explained by the residuals. The F value for a given relationship is compared to a table of

critical values for the F distribution for the appropriate number of regression ($n_1$) and residual ($n_2$) degrees of freedom. If the F value for the regression is found to be greater than the critical F value at a given confidence level, the regression is significant at this confidence level.

Table 2. Sample ANOVA Table

| Source | df† | SS†† | MS††† | F ratio | a§ |
|---|---|---|---|---|---|
| Regression | 1 | 6.326 | 6.326 | 6.569 | <0.01 |
| Residual | | 22 | 21.192 | 0.963 = $s^2$ | |
| Total, corrected SS | | 23 | 27.518 | | |

† degrees of freedom
†† sum of squares
††† mean square = SS/df
§ probability of significance test

Table 2 can be used to calculate the coefficient of determination, $r^2$:

$$r^2 = SS/\text{Total, corrected SS} \quad (7)$$

a value which indicates the proportion of the total variation about the mean regression line that is described by the regression equation. Table 2 data also allows the calculation of the standard deviation of the regression, s, which is the square root of the residual mean square. This value of s can be interpreted as the average residual, or average precision of the predicted value generated from the regression equation.

Student's t-test for the significance of variables

The coefficients estimated from the regression relationship are not known exactly, as each regression coefficient has a corresponding standard error. The Student's t value is the ratio of a regression coefficient value to its standard error, and indicates whether the coefficient value is significantly different from zero, i.e., whether it can be considered for inclusion in the

regression equation. The t value for a given coefficient is compared to values in a standard t table for n - 1 degrees of freedom, where n = the number of observations, to determine the probability level for the hypothesis that the coefficient is not statistically different from zero. If the probability from the standard t table is less than a specified probability level, 0.05 in this study, then the variable is identified as significant, and is included in the regression equation.

Precision of the predicted value

As described above, the standard deviation of the regression relationship, s, is a measure of the average precision of predicted values. However, as indicated by Draper and Smith (1981), the precision of of the predicted values depends on the values of the independent variables. These authors indicate that the s value underestimates the uncertainty associated with the predicted value, and that a better measure of the precision of $Y_i'$ at given values of the independent variable $X_o$ is given by the matrix equation:

$$Y_o = Y_o' \pm s\sqrt{1 + X_o' (X'X)^{-1} X_o} \qquad (8)$$

where $Y_o$ is the predicted value at a given value of X with the precision described as the product of s times the term under the radical, X is the matrix of independent variable values, and $X_o$ is the vector of independent variables at which predicted values of $Y_o$ are desired. As the number of observations increases, i.e., $n_{obs} > 500$, the term under the radical approaches one, and Equation 7 simplifies to the standard form:

$$Y_o = Y_o' \pm s \qquad (9)$$

Preliminary Results from MCI-Property Relationships

The relationship between estimated and experimental log Kow is illustrated in Figure 7.

The estimated log $K_{ow}$ values were calculated using the universal MCI-log $K_{ow}$ regression model. A significant improvement is observed when the four general class MCI-log $K_{ow}$ regression models (aromatic polar, aromatic non-polar, non-aromatic non-polar, and non-aromatic polar) are used in combination as shown in Figure 8.

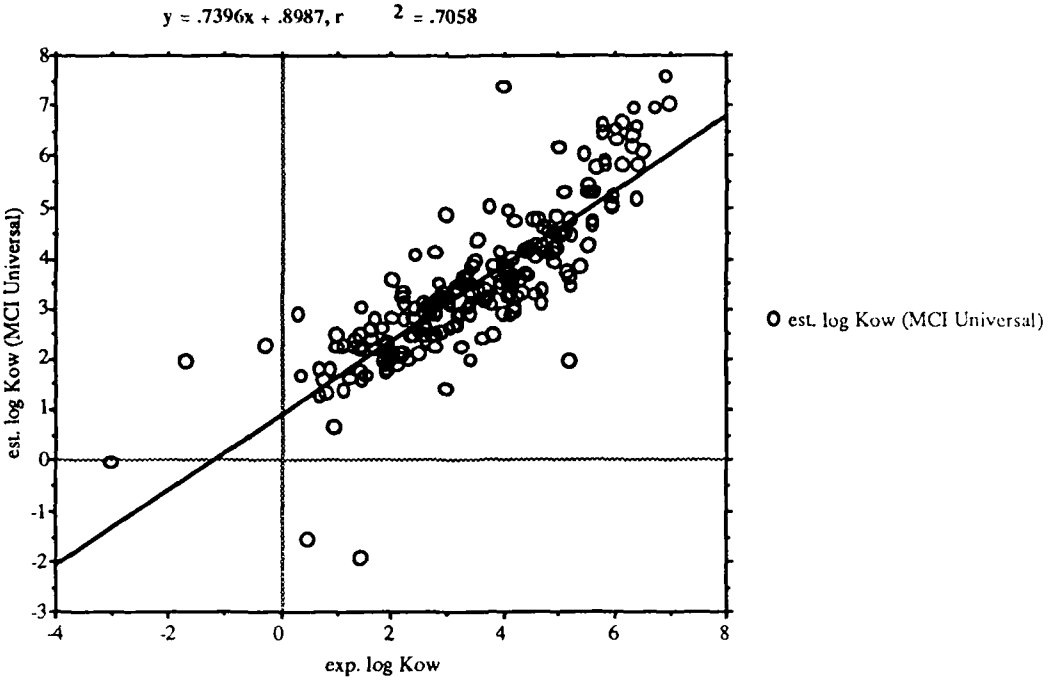



Figure 7. Experimental versus estimated (MCI Universal) log $K_{ow}$.

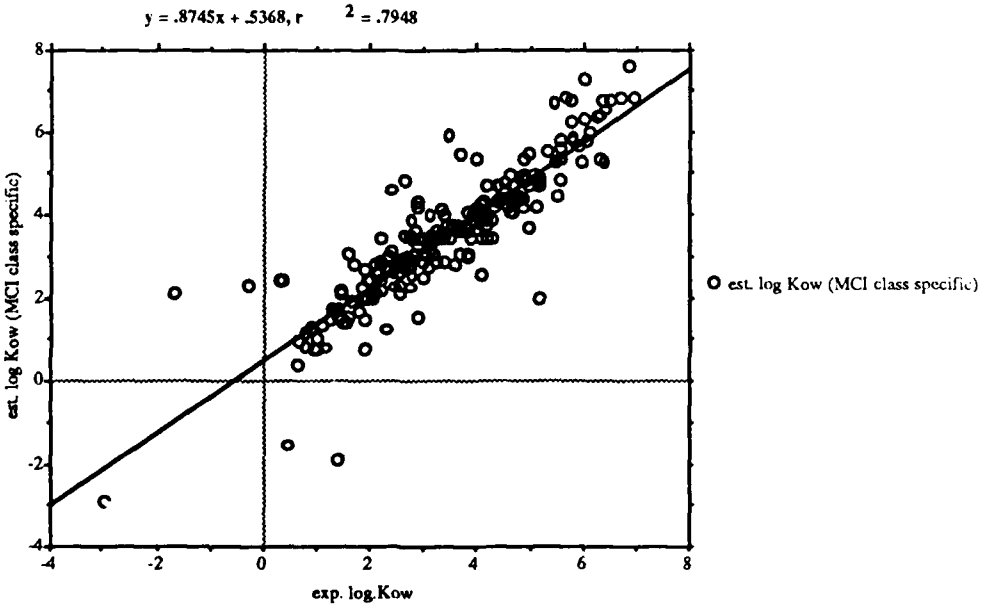



Figure 8. Experimental versus estimated (MCI four general equations) log $K_{ow}$.

For comparison, log $K_{ow}$ estimated using the computerized group contribution method (ClogP) of Leo and Hansch are plotted against experimental log $K_{ow}$ values for the same compounds used in Figures 7 and 8 (Figure 9). A comparison between properties estimated by PEP and those estimated by other established methods will be further explored in the project's second year.

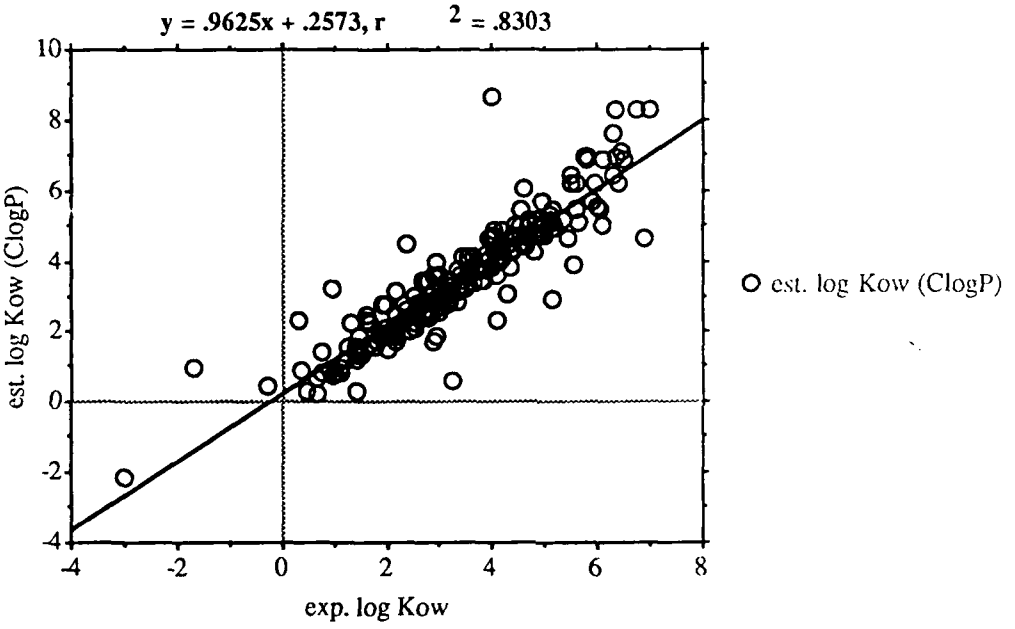



Figure 9. Experimental versus estimated (ClogP) log Kow.

## UNIFAC MODULE

### Overview

As discussed in the background and significance section UNIFAC-derived activity coefficients have been used to estimate values for aqueous solubility and octanol/water partition coefficients. However, since the UNIFAC approach is a solution of groups method, its operation is different than the MCI or property-property modules which are based on regression analysis.

Upon entering the UNIFAC module the user must first choose the property to be estimated. Currently, the UNIFAC module allows for the estimation of aqueous solubility and octanol/water partition coefficients. We have also included several additional properties which are of interest in environmental fate modeling: solubility in water/organic cosolvent systems, solubility in organic solvents, and oil/water partition coefficients. Only the aqueous solubility and octanol/water partition coefficient options are currently functional.

Upon entering the UNIFAC module the user must first choose the property to be estimated. After choosing the property, the required structural information must be input using one of three methods: manual, connection table (option not currently implemented), or SMILES string. The manual input method requires the user to dissect the molecule into its appropriate UNIFAC functional groups. A table of UNIFAC structural groups is provided for this purpose. If either the SMILES or connection table option is used the molecule is automatically dissected into its UNIFAC structural groups. Once the groups are selected the activity coefficent(s) can be calculated. The UNIFAC groups are displayed so that the user can verify their correctness. Figure 10 illustrates the overall operation of the UNIFAC module.

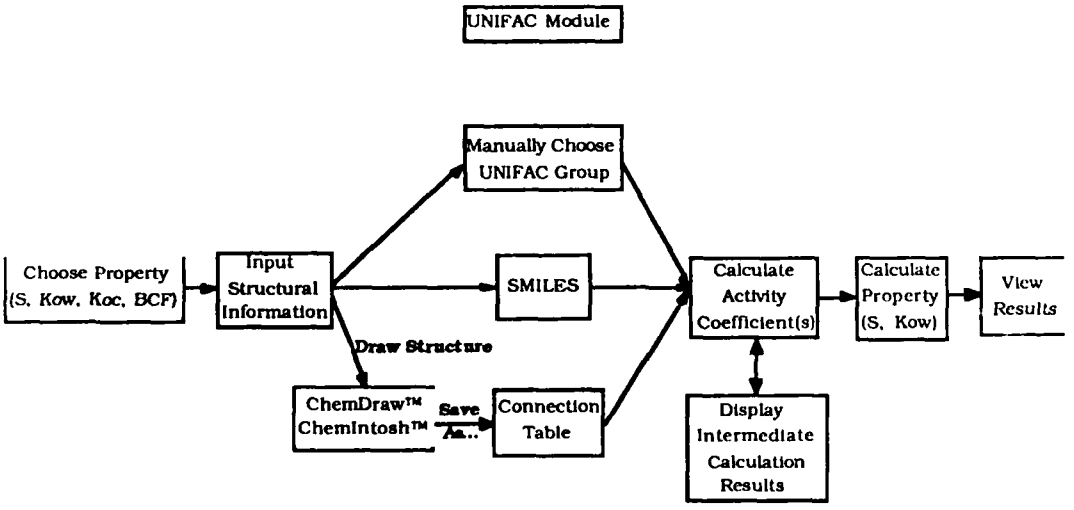



Figure 10. Flow chart depicting operation of UNIFAC module.

Calculation of UNIFAC derived activity coefficients

UNIFAC (UNIQUAC Functional Group Activity Coefficient) is a solution of groups method for calculating activity coefficients which requires only the structure of the compound as input. The calculated activity coefficients can be then used to estimate aqueous solubility, octanol/water partition coefficients and related properties such as solubility in organic solvents, solubility in mixed water organic solvents, and oil/water partition coefficients.

In this technique, UNIFAC calculates activity coefficients by dividing them into two parts, a combinatorial part which reflects the size and shape of the molecule present and a residual portion which depends on the functional group interactions:

$$\ln \gamma_i = \ln \gamma^C + \ln \gamma^R \tag{10}$$

where $\gamma_i$ is the activity coefficient for the ith molecular component in the mixture. The superscripts refer to the combinatorial (c) and residual parts (R)

Various parameters, such as van der Waals group volumes and surface areas and group interaction parameters, are input into a series of equations from which the combinatorial and residual parts are calculated. Values for the group parameters have been tabulated and can be found in the literature (Fredenslund et al., 1977 and Gmehling, 1982). The group parameters of Gmehling (1982) are used in the PEP UNIFAC module.

Estimation of aqueous solubility and octanol/water partition coefficients for UNIFAC derived activity coefficients

UNIFAC derived activity coefficients can be used to calculate aqueous solubility. The aqueous solubility (S) of an organic liquid can be approximated as follows:

$$S = 55.5\ a_{org}/\gamma_{aq} \tag{11}$$

where $a_{org}$ is the activity of the liquid in the organic phase and $\gamma_{aq}$ is the activity coefficient in the aqueous phase. For hydrophobic compounds, $a_{org}$ approximately one, since water does

not significantly dissolve in the organic phase. Thus the aqueous solubility of an organic liquid can be estimated from the following expression:

$$S = 55/(\gamma_{aq})_{UNIFAC} \quad (12)$$

where $(\gamma_{aq})_{UNIFAC}$ is the UNIFA C-derived activity coefficient at infinite dilution. For solids, the solubilities must be corrected to those of the corresponding supercooled liquids using the following expression (Yalkowski et al.,1980):

$$\log S_{supercooled\ liquid} = \log S_{solid} = 0.01(MP - 25) \quad (13)$$

where MP is the compound's melting point in °C.

Property/property correlation module

The property property correlation module is not fully implemented as the current time. Only a limited number of property/property correlations have been developed using the data collected in this study. A prototype of the module has been developed using property-property correlations published in the literature along with several general relationships developed in the first year of this project. The property-property module should be fully implemented with several months. A flow chart describing the operation of this prototype module is presented in Figure 11.

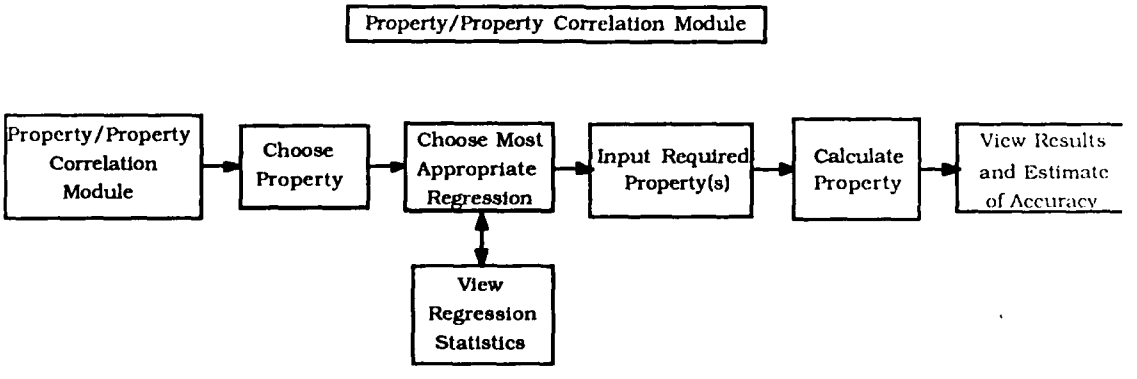



Figure 11. Flow chart depicting operation of UNIFAC module

Upon entering this module the user must first choose the property to be estimated. Once this is done the program then displays the available regression models along the required input

properties. At this time the user has the option of viewing the regression model statistics. Once the user chooses the most appropriate regression model and inputs the required information, the property can be estimated. The results of the property estimation can then be displayed along with an estimation of its accuracy based on the statistical approach described in the MCI section.

A summary of the property-property correlations developed in the first year of study are presented in Appendix B. It is anticipated that the module will be completed in several months.

## SUMMARY OF FIRST YEAR ACCOMPLISHMENTS

Chemical property data were compiled for over 700 compounds from a variety of literature sources and computerized databases. Only experimentally measured properties were used. Using this information, a chemical property database was created and used for developing MCI-property and property-property relationships which were then incorporated into a prototype microcomputer based property estimation program, referred to as PEP.

The property estimation program, PEP, is a decision support system, developed using HyperCard™ software, which utilizes MCI-property, property-property, and UNIFAC modules to provide the user with several approaches to estimate physical properties . The current version of PEP provides estimates of the following properties: aqueous solubility, octanol/water partition coefficients, vapor pressure, Henry's Law constants, organic carbon based soil sorption coefficients, and bioconcentration factors.

At the time of this report, the three property estimation modules are implemented, the MCI module is approximately 90% complete while the UNIFAC and property-property modules are approximately 50% complete.

## SECOND YEAR OBJECTIVES

In the original proposal, the major focus of the second year was to port the program developed on the Macintosh over to an MS-DOS platform. After reviewing the results obtained

in the first year, the investigators feel that further development and refining of the structural property relationships and linking PEP to an environmental fate model would be more important than implementing PEP on a different computer platform. Thus, we propose to accomplish the following tasks in the second year of the project:

1. Complete UNIFAC and Property-property modules.
2. Refine MCI-property relationships by examining the relationship between MCI and properties such as total molecular surface area, polarizability, and dipole moment.
3. Continue to develop chemical property database for current compounds and expand database by adding biotic and abiotic transformation related properties such as biodegradation, hydrolysis, and photolysis rates.
4. Investigate the relationship between MCI and transformation rates.
5. Link PEP with VIP (Vadose Zone Interactive Processes) to demonstrate the utility of combining a property estimation program with an environmental fate model.
6. If time permits, attempt to port PEP and associated database over to a MS-DOS platform.

## MISCELLANEOUS PUBLICATIONS

Frazier, J.D. 1990. Estimation of Chemical/Physical Properties Using Molecular Connectivity Indices for Application in Modeling the Environmental Fate of Organic Compounds. M.S., Utah State University.

## LIST OF PAPERS PRESENTED AT PROFESSIONAL MEETINGS

Doucette, W.J., H. Fugate, J.D. Frazier, and D.A. Bahnick. 1989. Structure Property Relationships Utilizing Molecular Connectivity Indices and Total Molecular Surface Areas for Environmental Fate Modeling. Presented at the International Chemical Congress of Pacific Basin Societies, December 21, Honolulu, Hawaii, USA

Frazier, J.D, and W.J. Doucette. 1989. Microcomputer Program Utilizing Molecular Connectivity Indices for Property Estimation. Presented at the International Chemical Congress of Pacific Basin Societies, December 21, Honolulu, Hawaii, USA

W.J. Doucette, D.J Denne, J.D. Frazier, R.R. Dupont and D.K. Stevens. 1990. Microcomputer Program Utilizing Molecular Connectivity Indices, Unifac-Derived Activity

Coefficients, And Property-Property Correlations For The Estimation Of Physical Properties. Presented at the Fourth International Workshop on QSAR in Environmental Toxicology, September 16-20, Veldhoven, The Netherlands.

## LIST OF GRADUATE STUDENTS ASSOCIATED WITH THE RESEARCH EFFORT

Doug Denne, M.S. expected December 1990
Mark Holt, M.S. expected December 1991
Joe Frazier, M.S. 1990

## REFERENCES


Anscombe, F. J., and J. W. Tukey. 1963. The examination and analysis of residuals. Technometrics, 5:141-160.

Arbuckle, W. B. 1983. Using UNIFAC to Calculate Aqueous Solubilities. Environ. Sci. Technol., 20, 1060-1064.

Bahnick, D.A. and W. J. Doucette. 1988. Use of Molecular Connectivity Indices to Estimate Soil Sorption Coefficients for Organic Chemicals. Chemosphere. 17:1703-1715.

Briggs, G. G. 1981. Theoretical and Experimental Relationships between Soil Adsorption, Octanol-Water partition Coefficients, Water Solubilities, Bioconcentration Factors and the Parachor. J. Agric. Food Chem., 29, 1050-1059.

Chiou, C.T., V. H. Freed, D. W. Schmedding and R. L. Kohnert. 1977. Partition Coefficients and Bioaccumulation of Selected Organic Chemicals. Environ Sci. Technol. 11:475-478.

Doucette, W. J. and Andren, A.W. 1987. Correlation of Octanol/Water Partition Coefficients and Total Molecular Surface Area of Highly Hydrophobic Aromatic Compounds. Environ. Sci. Technol., 21, 821-824.

Doucette, W. J. and A. W. Andren. (1988). "Estimation of Octanol/Water Partition Coefficients: Evaluation of Six Methods for Highly Hydrophobic Aromatic Hydrocarbons." Chemosphere. 17(2): 334-359.

Draper, N. R., and H. Smith. 1981. Applied Regression Analysis. John Wiley and Sons, New York, NY. 709 pp.

Fredenslund, A.J., Gmehling and P. Rasmussen. 1977. Vapor-liquid equilibria using UNIFAC, a group-contribution method. Editor. Vapor-liquid equilibria using UNIFAC, a group-contribution method. Elsevier Scientific Publishing Co. New York, NY 213 pp.

Gerstl, A. and C. S. Helling. 1987. Evaluation of Molecular Connectivity as a predicitve method for the adsorption of pesticides by soils. J. Environ. Sci. Health. B22:55-69.

Gmehling, J. P. Rasmussen and A. Fredenslund. 1982. Vapor-liquid equilibria by group contribution. Revision and Extension. 2. Ind. Eng. Chem. Process Des. Dev. 21:118-127.

Hansch, C. and Leo, A.J. 1979. Substituent Constants for Correlation Analysis in Chemistry and Biology. John Wiley, New York, NY. 339pp.

Kenaga, E.E., and C.A. I. Goring. 1980. Relationship Between Water Solubility, Soil Sorption, Octanol-Water Partitioning, and Concentration of Chemicals in Biota. American Society for Testing and Materials : 78-115.

Kier, L.B. 1980. Molecular connectivity as a description of structure for SAR analysis in Physical Chemical Properties of Drugs. Editor. Molecular connectivity as a description of structure for SAR analysis in Physical Chemical Properties of Drugs. Yalkowsky, S. H., A. A. Sikula, and S. C. Valvani, eds. Marcel Dekker, Inc. 331pp. New York, NY

Kier, L.B., Hall and L. H. 1986. Molecular Connectivity in Structure-Activity Analysis. Editor. Molecular Connectivity in Structure-Activity Analysis. Research Studies Press, LTD., England, and John Wiley & Sons, Inc., New York. Letchworth, Hertfordshire, England.

Kier, L.B. and L. H. Hall. 1976. Molecular connectivity in chemistry and drug research. in Medicinal Chemistry a Series of Monographs. Editor. Molecular connectivity in chemistry and drug research. in Medicinal Chemistry a Series of Monographs. DeSteven, G. ed. Academic Press. N. Y.

Koneman, H. (1980). "Structure-activity relationships and additivity in fish toxicology of environmental pollutants." Ecotoxicology and Environmental Safety. 4: 415-421.

Leo, A.J. 1985. CLOGP-3.3 computer program, MedChem Project, Pomona College, Clairemont, CA.

Lowell H. Hall, Lemont B. Kier, and Wallace J. Murray. 1974. Molecular Connectivity II: Relationship to water solubility and Boiling point. Journal of Pharmaceutical Sciences, 64, 1974-1977.

Lyman, W.J.. 1985. Estimation of Physical Properties in Environmental Exposure from Chemicals. Editor. Estimation of Physical Properties in Environmental Exposure from Chemicals. W. B. Neely, and G. E. Blau, Eds. CRC Press. Boca Raton, FLp14-44 pp

Lyman, W.J., W. F. Reehl and D. H. Rosenblatt. 1982. Handbook of Chemical Property Estimation Methods. Environmental Behavior of Organic Compounds. Editor. Handbook of Chemical Property Estimation Methods. Environmental Behavior of Organic Compounds. McGraw-Hill Inc. New York, NY. 860 pp

Mackay, D., R. Mascarenhas and W. Y. Shiu. (1980). "Aqueous Solubility of Polychlorinated Biphenyls." Chemosphere. 9: 257-264.

Mackay, D., A. Bobra, D. W. Chan, and W. Y. Shiu. 1982.Vapor Pressure Correlations for Low Volatility Environmental Chemicals. Environ. Sci. Technol., 16, 645-649

Randie, M. (1975). "On Characterization of Molecular Branching." J. Amer. Chem. Soc. 97(23): 6609-6615.

Sabljic, A. 1984. Prediction of the Nature and Strength of Soil Sorption of Organic Pollutants by Molecular Topology. J. Agric. Food Chem., 32, 243-246.

Sabljic, A. 1987. On the Prediction of Soil Sorption Coefficients of Organic Pollutants from Molecular Structure: Application of Molecular Topology Model. Environ. Sci. Technol., 21, 358-366.

# APPENDIX A

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: S Universal

## Regression Results

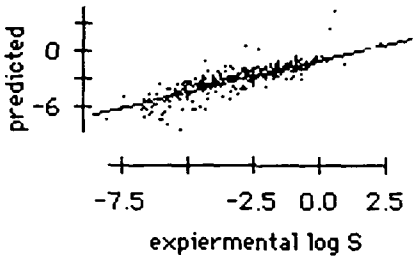
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.3917 | 0.1376 | 2.85 |
| vp1 | -.9257 | 0.0316 | -29.3 |
| Δvp1 | 1.8251 | 0.1047 | 17.4 |

## Analysis of Variance Table

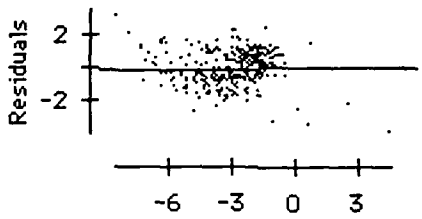
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 889.176 | 2 | 445 | 446 |
| Residual | 360.920 | 362 | 0.997 | |
| Total | 1250.096 | 364 | 3.4343 | |

$r^2$ = 71.1% $n_{obs}$ = 365 s = 0.9985

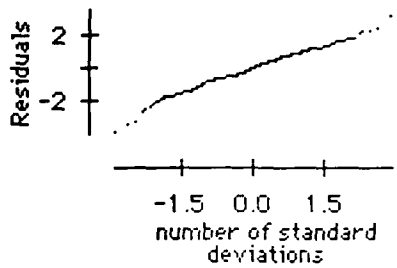
**Predicted vs. Exp.**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

# STATISTICS Class: S Universal -ionizable

## Regression Results

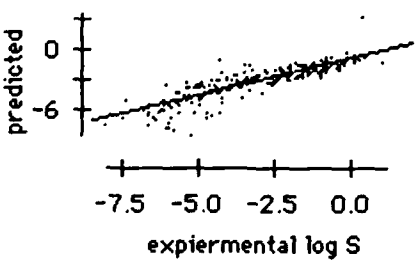
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.437 | 0.1379 | 3.17 |
| vp1 | -.9748 | 0.0323 | -30.2 |
| Δvp1 | 2.129 | 0.1282 | 16.6 |

## Analysis of Variance Table

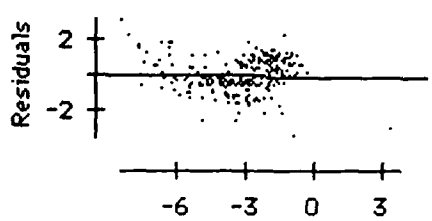
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 793.809 | 2 | 397 | 462 |
| Residual | 245.514 | 286 | 0.8584 | |
| Total | 1039.323 | 288 | 3.6088 | |

$r^2$ = 76.4% $n_{obs}$ = 289 s = 0.9265

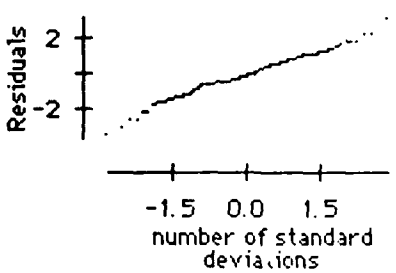
**Predicted vs. Exp.**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: non-polar, non-cyclic

## Regression Results

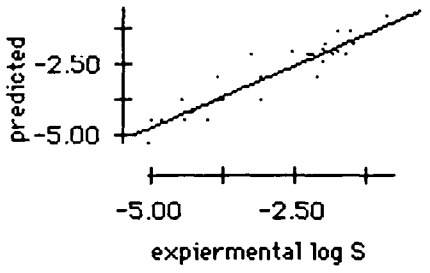
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 2.0477 | 0.3149 | 6.5 |
| np1 | -2.0254 | 0.1515 | -13.4 |
| np3 | 0.9608 | 0.1625 | 5.91 |

## Analysis of Variance Table

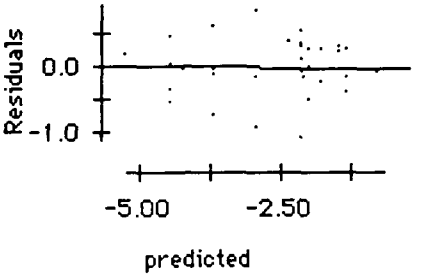
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 52.8732 | 2 | 26.44 | 141 |
| Residual | 6.00701 | 32 | 0.1877 | |
| Total | 58.88021 | 34 | 1.7318 | |

$r^2$ = 89.8% $n_{obs}$ = 35 s = 0.4333

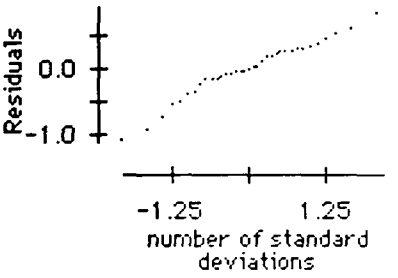
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

# STATISTICS Class: S Oxygenated aliphatics

## Regression Results

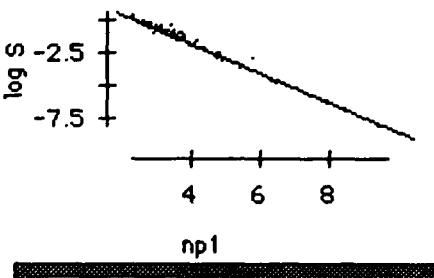
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 2.9022 | 0.074 | 39.2 |
| np1 | -1.1561 | 0.0178 | -65.1 |

## Analysis of Variance Table

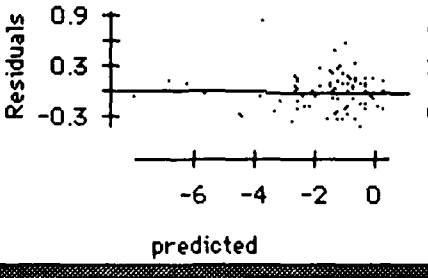
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 196.290 | 1 | 196 | 4240 |
| Residual | 3.65699 | 79 | 0.0463 | |
| Total | 199.9469 | 80 | 2.4993 | |

$r^2$ = 98.2% $n_{obs}$ = 81 s = 0.2152

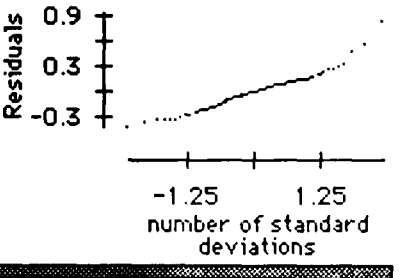
Log S vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: S non-polar, Aromatic

## Regression Results

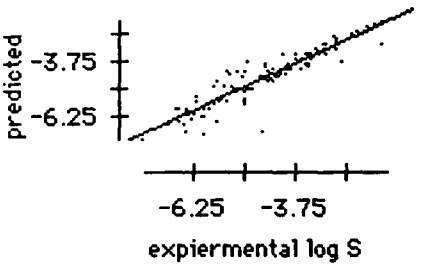
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.2478 | 0.2122 | 1.17 |
| vp1 | -1.1114 | 0.0579 | -19.2 |
| vp6 | 0.6617 | 0.1107 | 5.98 |

## Analysis of Variance Table

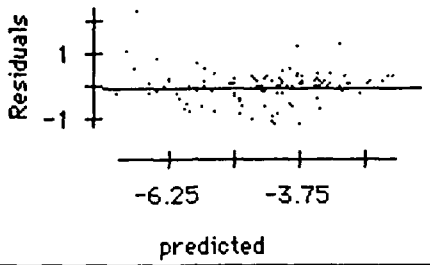
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 169.662 | 2 | 84.8 | 349 |
| Residual | 26.9593 | 111 | 0.2429 | |
| Total | 196.6213 | 113 | 1.74 | |

$r^2$ = 86.3% $n_{obs}$ = 114 s = 0.4928

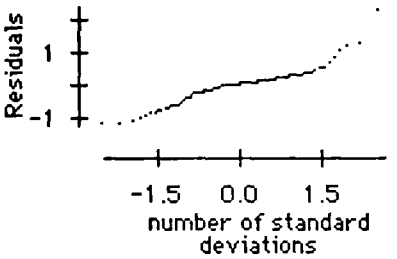
**Predicted vs. Exp.**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

# STATISTICS Class: S PCBs

## Regression Results

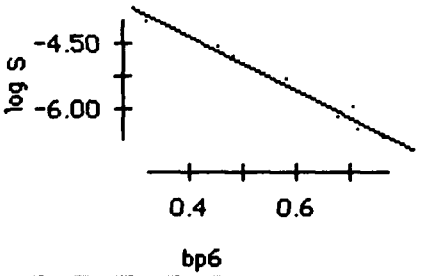
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -1.8528 | 0.2204 | -8.41 |
| bp6 | -6.2017 | 0.3654 | -17 |

## Analysis of Variance Table

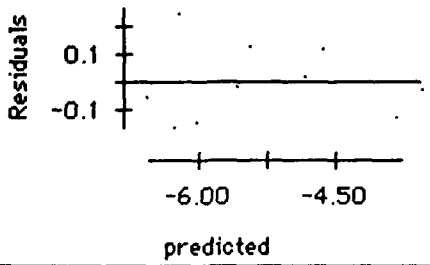
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 6.50320 | 1 | 6.5032 | 288 |
| Residual | 0.158026 | 7 | 0.0226 | |
| Total | 6.661226 | 8 | 0.8327 | |

$r^2$ = 97.6% $n_{obs}$ = 9 s = 0.1503

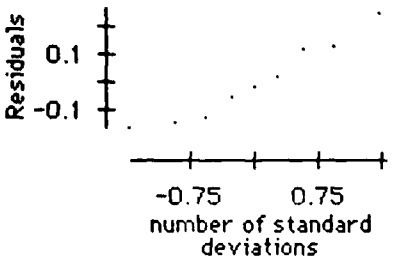
**Log S vs. MCI**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: S ureas

## Regression Results

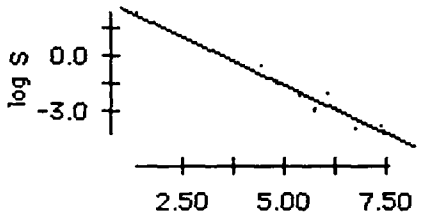
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 3.758 | 0.4631 | 8.12 |
| bp1 | -1.0747 | 0.0808 | -13.3 |

## Analysis of Variance Table

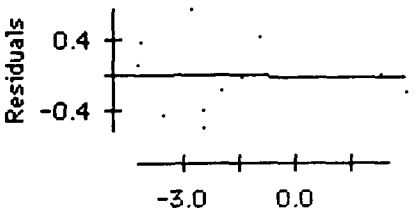
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 32.5444 | 1 | 32.54 | 177 |
| Residual | 1.65604 | 9 | 0.184 | |
| Total | 34.20044 | 10 | 3.42 | |

$r^2$ = 95.2% $n_{obs}$ = 11 s = 0.4290

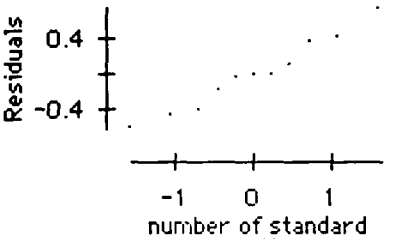
Log S vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

# STATISTICS Class: S Anilines

## Regression Results

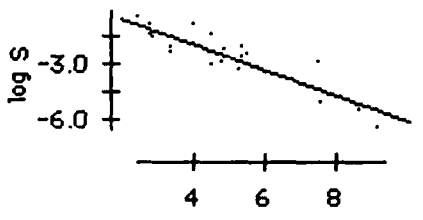
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.9892 | 0.401 | 2.47 |
| bp1 | -.7157 | 0.0734 | -9.75 |

## Analysis of Variance Table

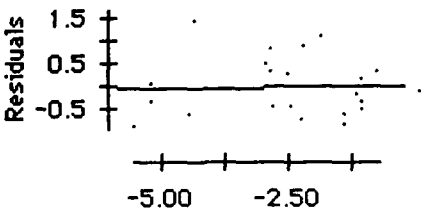
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 45.5658 | 1 | 45.57 | 95.0 |
| Residual | 9.10898 | 19 | 0.4794 | |
| Total | 54.67478 | 20 | 2.7337 | |

$r^2$ = 83.3% $n_{obs}$ = 21 s = 0.6924

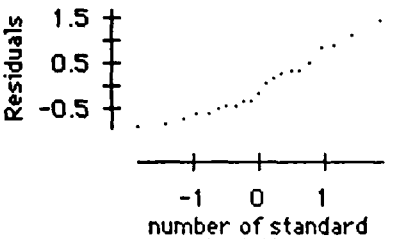
Log S vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: S Alcohols

## Regression Results

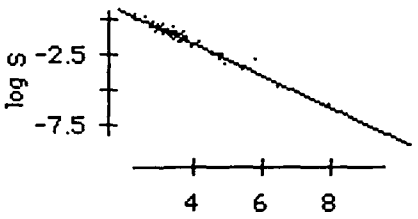
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 2.8524 | 0.0839 | 34 |
| bp1 | -1.1446 | 0.0201 | -57 |

## Analysis of Variance Table

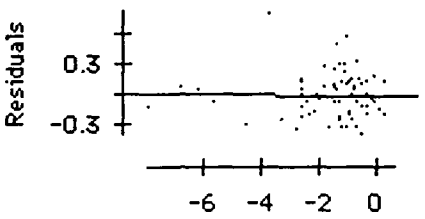
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 167.948 | 1 | 168 | 3248 |
| Residual | 3.41245 | 66 | 0.0517 | |
| Total | 171.3604 | 67 | 2.5576 | |

$r^2$ = 98.0% $n_{obs}$ = 68 s = 0.2274

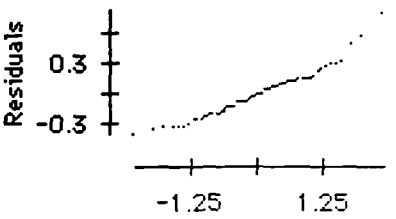
Log S vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

# STATISTICS Class: S Halogenated benzenes

## Regression Results

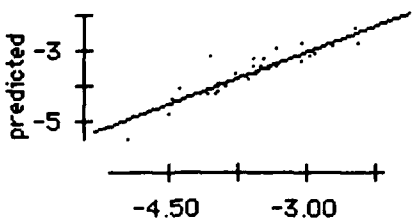
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -.7166 | 0.2099 | -3.41 |
| nc3 | -.8638 | 0.3031 | -2.85 |
| vp0 | -.3291 | 0.0405 | -8.12 |

## Analysis of Variance Table

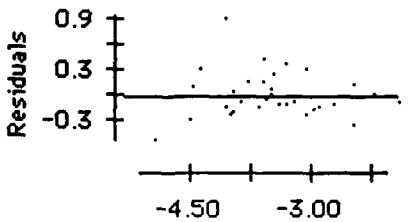
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 15.5443 | 2 | 7.772 | 102 |
| Residual | 2.21780 | 29 | 0.0765 | |
| Total | 17.7621 | 31 | 0.573 | |

$r^2$ = 87.5% $n_{obs}$ = 32 s = 0.2765

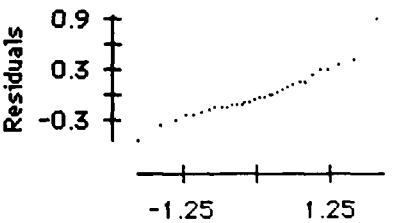
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: S PAHs

## Regression Results

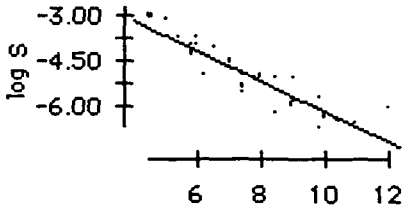
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -1.1353 | 0.2894 | -3.92 |
| np1 | -.5011 | 0.0373 | -13.4 |

## Analysis of Variance Table

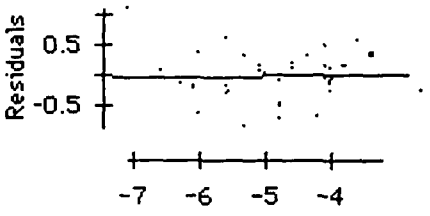
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 30.5352 | 1 | 30.54 | 181 |
| Residual | 5.74826 | 34 | 0.1691 | |
| Total | 36.28346 | 35 | 1.0367 | |

$r^2$ = 84.2% $n_{obs}$ = 36 s = 0.4112

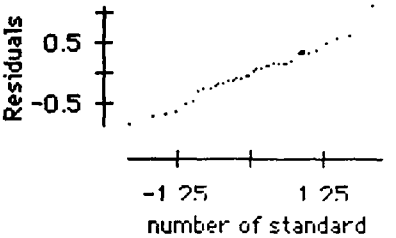
**Log S vs. MCI**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

# STATISTICS Class: S Polar Aromatics

## Regression Results

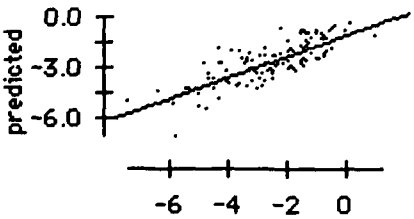
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 1.4051 | 0.3001 | 4.68 |
| vp0 | -.4987 | 0.0396 | -12.6 |
| Δvp1 | 0.5967 | 0.2204 | 2.71 |

## Analysis of Variance Table

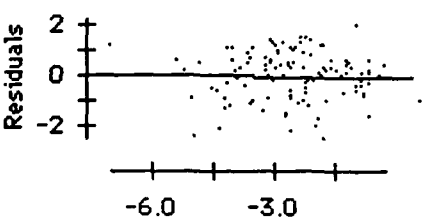
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 174.537 | 2 | 87.3 | 97.7 |
| Residual | 96.5006 | 108 | 0.8935 | |
| Total | 271.0376 | 110 | 2.464 | |

$r^2$ = 64.4% $n_{obs}$ = 111 s = 0.9453

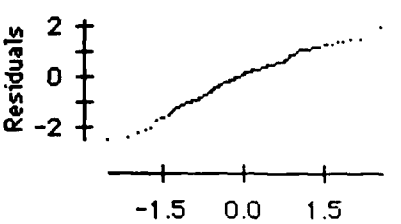
**Predicted vs. Exp.**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: Pv PCBs

## Regression Results

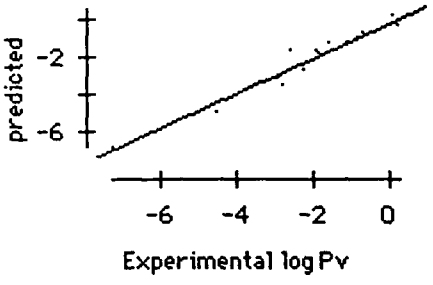
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 9.4791 | 1.815 | 5.22 |
| np3 | -2.4284 | 0.4481 | -5.42 |
| nc5 | 5.8143 | 2.15 | 2.7 |

## Analysis of Variance Table

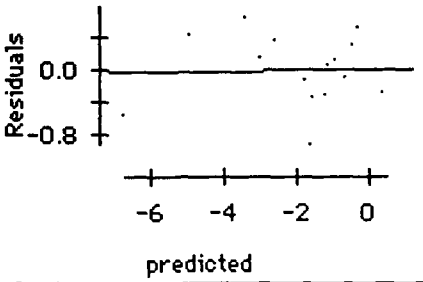
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 49.8285 | 2 | 24.91 | 115 |
| Residual | 2.60424 | 12 | 0.217 | |
| Total | 52.43274 | 14 | 3.7452 | |

$r^2$ = 95.0% $n_{obs}$ = 15 s = 0.4659

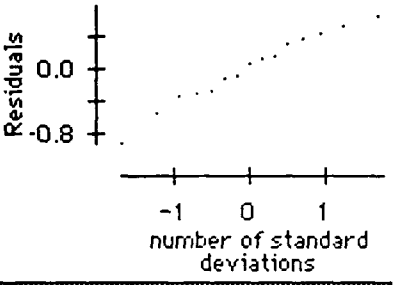
**Predicted vs. Exp.**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

# STATISTICS Class: Pv Universal

## Regression Results

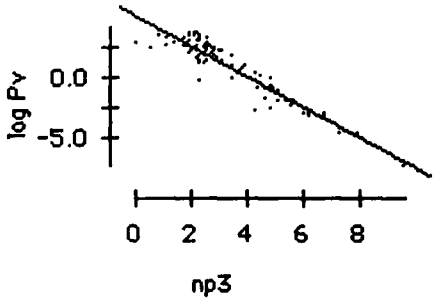
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 5.2611 | 0.1801 | 29.2 |
| np3 | -1.2746 | 0.0491 | -25.9 |

## Analysis of Variance Table

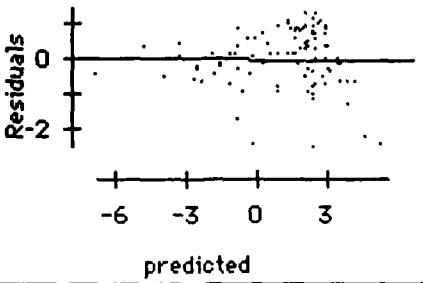
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 453.996 | 1 | 454 | 672 |
| Residual | 62.7848 | 93 | 0.675 | |
| Total | 516.7808 | 94 | 5.4977 | |

$r^2$ = 87.9% $n_{obs}$ = 95 s = 0.8216

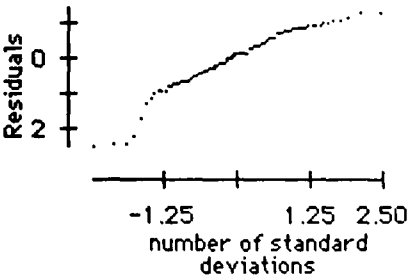
**Log Pv vs. MCI**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: Pv PAHs

### Regression Results

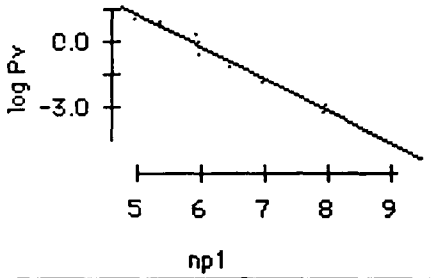
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 8.7168 | 0.4656 | 18.7 |
| np1 | -1.4892 | 0.0702 | -21.2 |

### Analysis of Variance Table

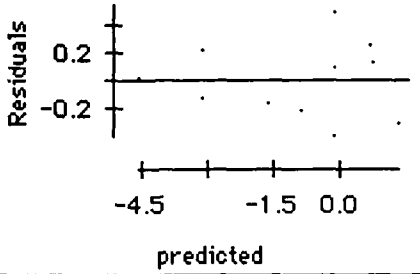
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 35.5950 | 1 | 35.6 | 450 |
| Residual | 0.712338 | 9 | 0.0791 | |
| Total | 36.30733 | 10 | 3.6307 | |

$r^2$ = 98.0% $n_{obs}$ = 11 s = 0.2813

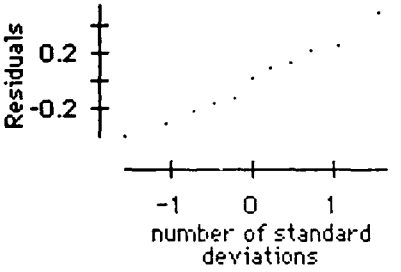
Log Pv vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

## STATISTICS Class: Pv Halogenated Aromatics

### Regression Results

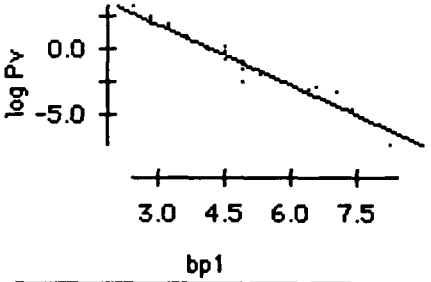
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 6.622 | 0.349 | 19 |
| bp1 | -1.5592 | 0.0701 | -22.2 |

### Analysis of Variance Table

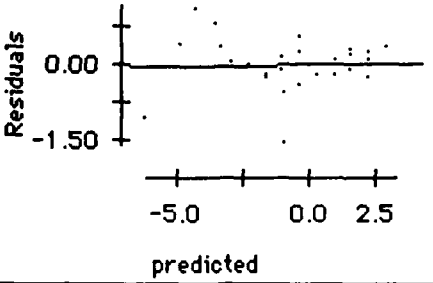
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 147.169 | 1 | 147 | 494 |
| Residual | 7.14780 | 24 | 0.2978 | |
| Total | 154.3168 | 25 | 6.1727 | |

$r^2$ = 95.4% $n_{obs}$ = 26 s = 0.5457

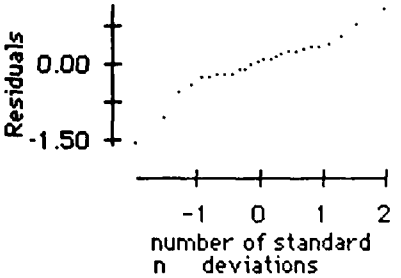
Log Pv vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: Kow Universal

### Regression Results

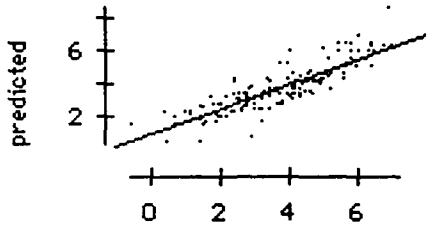
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 1.1527 | 0.1411 | 8.17 |
| vp0 | 0.3978 | 0.0175 | 22.7 |
| Δvp0 | -1.9843 | 0.1234 | -16.1 |

### Analysis of Variance Table

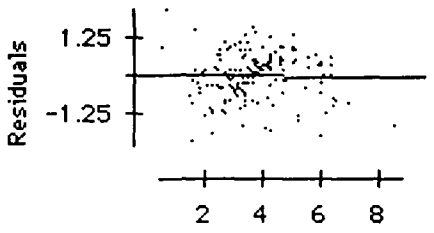
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 335.264 | 2 | 168 | 315 |
| Residual | 102.809 | 193 | 0.5327 | |
| Total | 438.073 | 195 | 2.2465 | |

$r^2$ = 76.5% $n_{obs}$ = 196 s = 0.7299

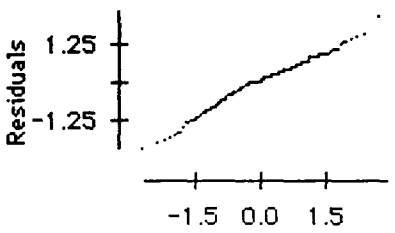
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

## STATISTICS Class: Kow ureas

### Regression Results

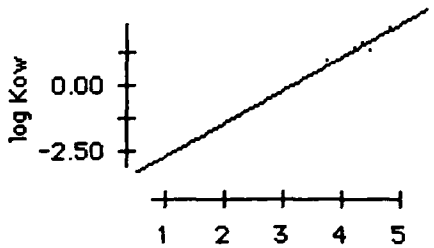
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -3.9333 | 0.2288 | -17.2 |
| vp1 | 1.2594 | 0.0561 | 22.5 |

### Analysis of Variance Table

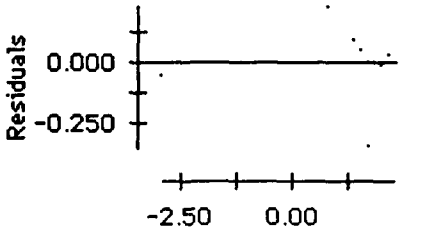
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 18.8299 | 1 | 18.83 | 504 |
| Residual | 0.186679 | 5 | 0.0373 | |
| Total | 19.01657 | 6 | 3.1694 | |

$r^2$ = 99.0% $n_{obs}$ = 7 s = 0.1932

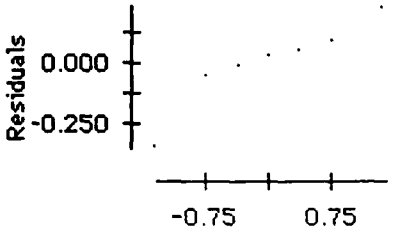
Log Kow vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: Kow Polar Aromatics

## Regression Results

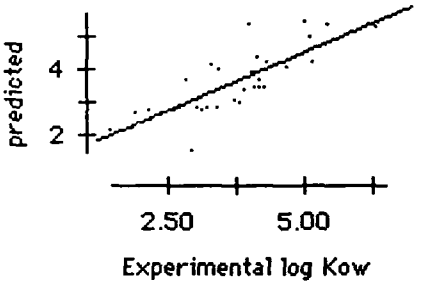
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 1.8347 | 0.2632 | 6.97 |
| vp0 | 0.2869 | 0.028 | 10.2 |
| Δvp0 | -1.4152 | 0.2617 | -5.41 |

## Analysis of Variance Table

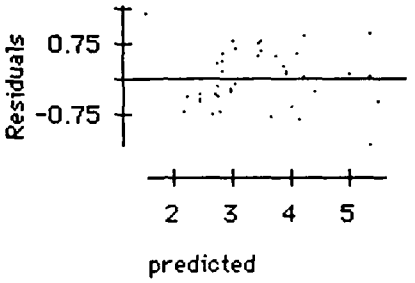
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 39.0336 | 2 | 19.52 | 52.9 |
| Residual | 14.7522 | 40 | 0.3688 | |
| Total | 53.7858 | 42 | 1.2806 | |

$r^2$ = 72.6% $n_{obs}$ = 43 s = 0.6073

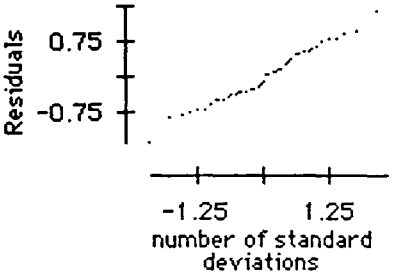
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

# STATISTICS Class: Kow PCBs

## Regression Results

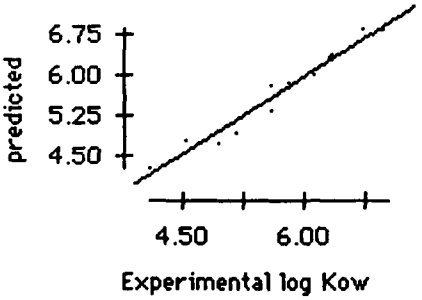
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 3.554 | 0.1888 | 18.8 |
| bp4 | -.9026 | 0.2567 | -3.52 |
| bp6 | 5.8313 | 0.6189 | 9.42 |

## Analysis of Variance Table

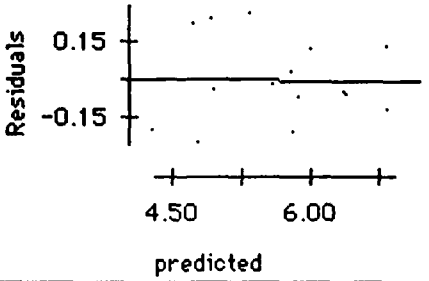
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 9.80627 | 2 | 4.9031 | 167 |
| Residual | 0.382424 | 13 | 0.0294 | |
| Total | 10.18869 | 15 | 0.6792 | |

$r^2$ = 96.2% $n_{obs}$ = 16 s = 0.1715

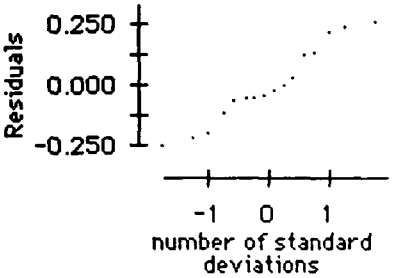
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: Kow PAH's

## Regression Results

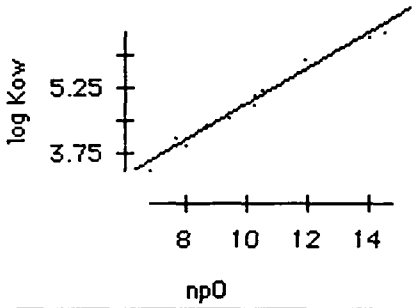
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.7812 | 0.218 | 3.58 |
| np0 | 0.4103 | 0.0206 | 19.9 |

## Analysis of Variance Table

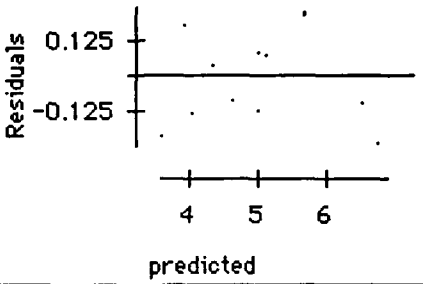
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 10.9031 | 1 | 10.90 | 395 |
| Residual | 0.303592 | 11 | 0.0276 | |
| Total | 11.20669 | 12 | 0.9339 | |

$r^2$ = 97.3% $n_{obs}$ = 13 s = 0.1661

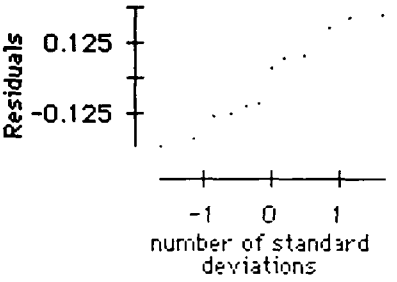
Log Kow vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

# STATISTICS Class: Kow nonpolar aromatic

## Regression Results

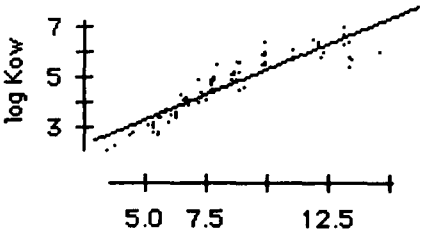
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 1.2033 | 0.1598 | 7.53 |
| vp0 | 0.4146 | 0.0187 | 22.2 |

## Analysis of Variance Table

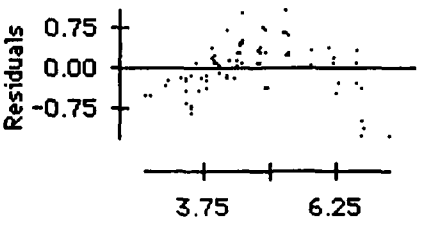
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 102.413 | 1 | 102 | 494 |
| Residual | 17.6377 | 85 | 0.2075 | |
| Total | 120.0507 | 86 | 1.3959 | |

$r^2$ = 85.3% $n_{obs}$ = 87 s = 0.4555

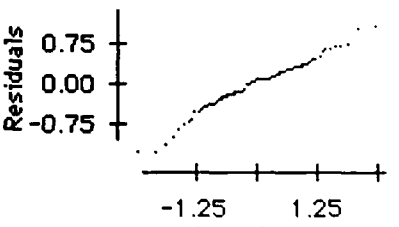
Log Kow vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: Kow noncyclic Oxygenated

### Regression Results

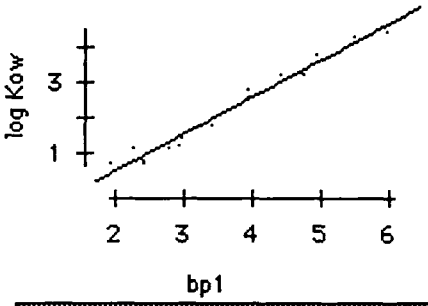
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -1.6181 | 0.1489 | -10.9 |
| bp1 | 1.0591 | 0.0397 | 26.7 |

### Analysis of Variance Table

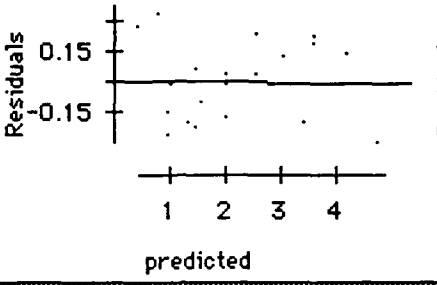
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 28.7287 | 1 | 28.73 | 713 |
| Residual | 0.725310 | 18 | 0.0403 | |
| Total | 29.45401 | 19 | 1.5502 | |

$r^2$ = 97.5% $n_{obs}$ = 20 s = 0.2007

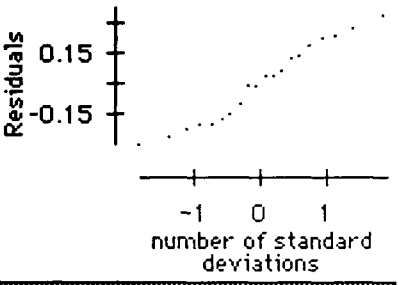
Log Kow vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

## STATISTICS Class: Kow noncyclic nonpolar

### Regression Results

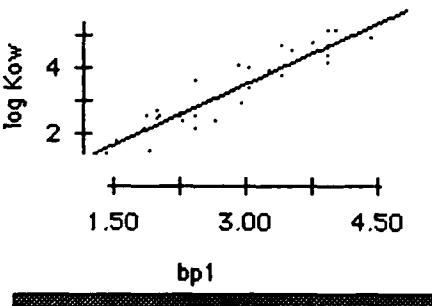
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -.2407 | 0.2418 | -.995 |
| bp1 | 1.2633 | 0.0842 | 15 |

### Analysis of Variance Table

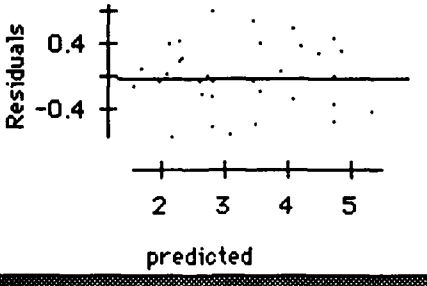
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 37.0139 | 1 | 37.01 | 225 |
| Residual | 5.26621 | 32 | 0.1646 | |
| Total | 42.2801 | 33 | 1.2812 | |

$r^2$ = 87.5% $n_{obs}$ = 34 s = 0.4057

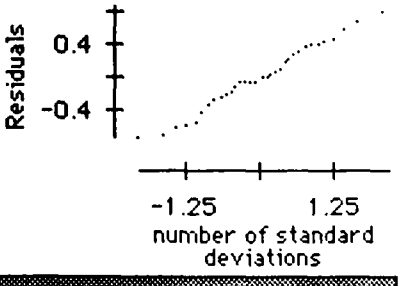
Log Kow vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: Kow anilines

## Regression Results

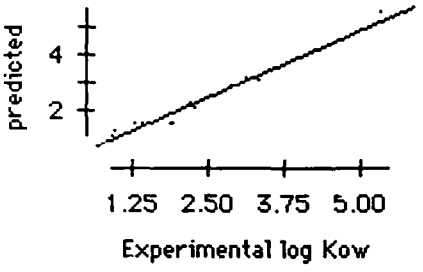
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -1.3184 | 0.2381 | -5.54 |
| bp1 | 1.0828 | 0.0878 | 12.3 |
| Δvp5 | -3.9803 | 0.6307 | -6.31 |

## Analysis of Variance Table

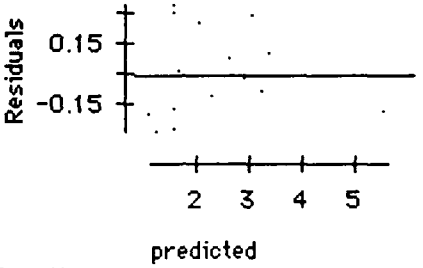
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 18.5293 | 2 | 9.265 | 172 |
| Residual | 0.646843 | 12 | 0.0539 | |
| Total | 19.17614 | 14 | 1.3697 | |

$r^2$ = 96.6% $n_{obs}$ = 15 s = 0.2322

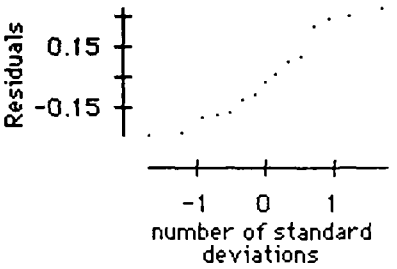
**Predicted vs. Exp.**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

# STATISTICS Class: Kow Alcohols

## Regression Results

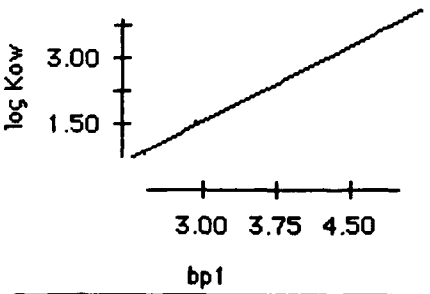
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -1.9613 | 0.1281 | -15.3 |
| bp1 | 1.1665 | 0.0354 | 32.9 |

## Analysis of Variance Table

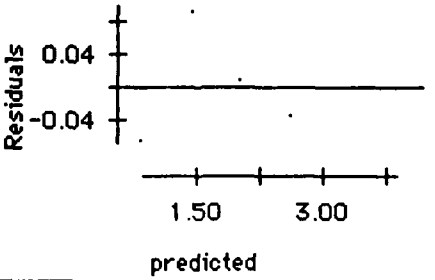
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 5.03456 | 1 | 5.0346 | 1085 |
| Residual | 0.013924 | 3 | 0.0046 | |
| Total | 5.048484 | 4 | 1.2621 | |

$r^2$ = 99.7% $n_{obs}$ = 5 s = 0.0681

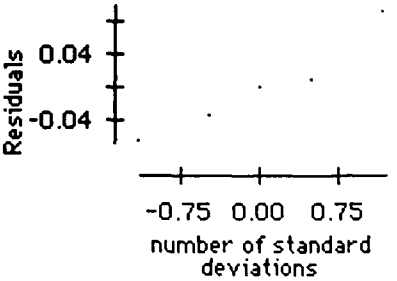
**Log Kow vs. MCI**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: Koc Universal

## Regression Results

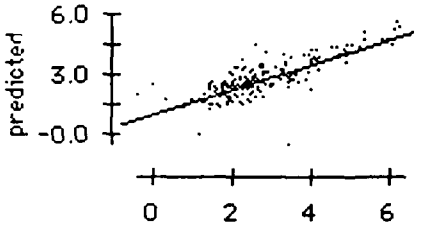
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.6111 | 0.1767 | 3.46 |
| np1 | 0.4647 | 0.0289 | 16.1 |
| Δvp0 | -1.2243 | 0.1176 | -10.4 |

## Analysis of Variance Table

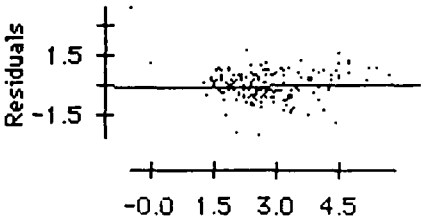
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 159.141 | 2 | 79.6 | 138 |
| Residual | 96.6154 | 168 | 0.5751 | |
| Total | 255.7564 | 170 | 1.5044 | |

$r^2$ = 62.2% $n_{obs}$ = 171 s = 0.7583

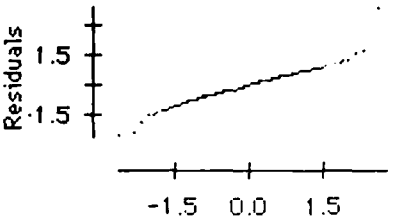
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

# STATISTICS Class: Koc Ureas 2

## Regression Results

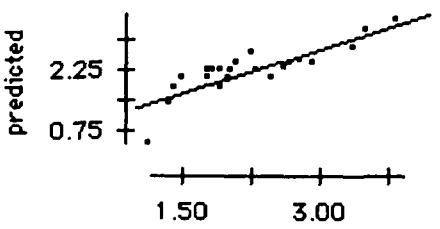
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -.4476 | 0.4502 | -.994 |
| Δvp1 | -.0235 | 0.4433 | -.053 |
| np0 | 0.2568 | 0.0385 | 6.68 |

## Analysis of Variance Table

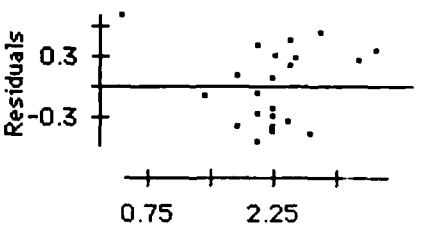
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 7.64836 | 2 | 3.8242 | 24.8 |
| Residual | 2.92537 | 19 | 0.154 | |
| Total | 10.57373 | 21 | 0.5035 | |

$r^2$ = 73.2% $n_{obs}$ = 22 s = 0.3924

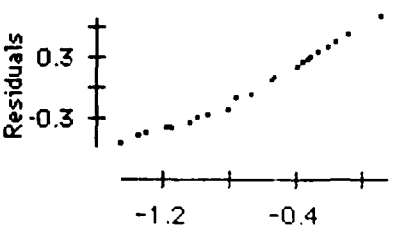
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: Koc Ureas

### Regression Results

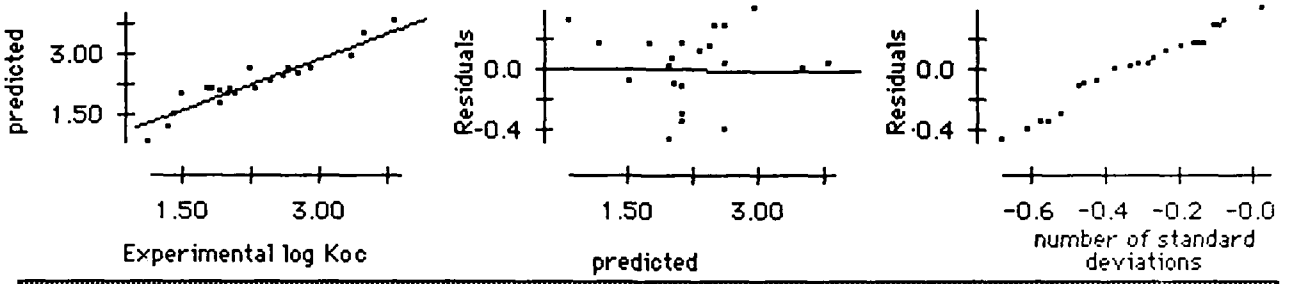
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.8342 | 0.1578 | 5.29 |
| np6 | 1.6588 | 0.148 | 11.2 |
| nch6 | -6.2556 | 1.515 | -4.13 |

### Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 9.29029 | 2 | 4.6451 | 68.8 |
| Residual | 1.28344 | 19 | 0.0675 | |
| Total | 10.5737 | 21 | 0.5035 | |

$r^2$ = 87.9% $n_{obs}$ = 22 s = 0.2599

Predicted vs. Exp. Residual vs. Predicted Residual vs. Prob.




Help List Return to PEP data base

## STATISTICS Class: Koc Triazines

### Regression Results

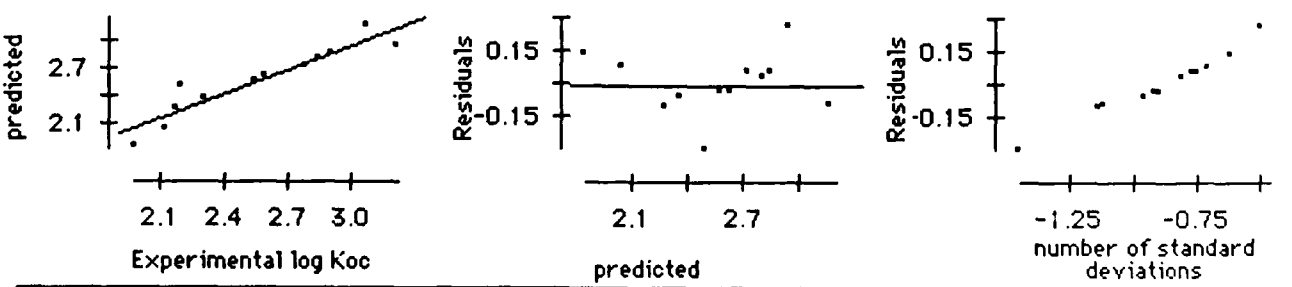
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -.5643 | 0.7221 | -.781 |
| bp1 | 0.6296 | 0.0919 | 6.85 |
| Δvp0 | -.9382 | 0.3763 | -2.49 |

### Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 1.60125 | 2 | 0.80062 | 32.6 |
| Residual | 0.221318 | 9 | 0.0246 | |
| Total | 1.8226 | 11 | 0.1657 | |

$r^2$ = 87.9% $n_{obs}$ = 12 s = 0.1568

Predicted vs. Exp. Residual vs. Predicted Residual vs. Prob.




Help List Return to PEP data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: Koc Polar, non-aromatic, non-cyclic

### Regression Results

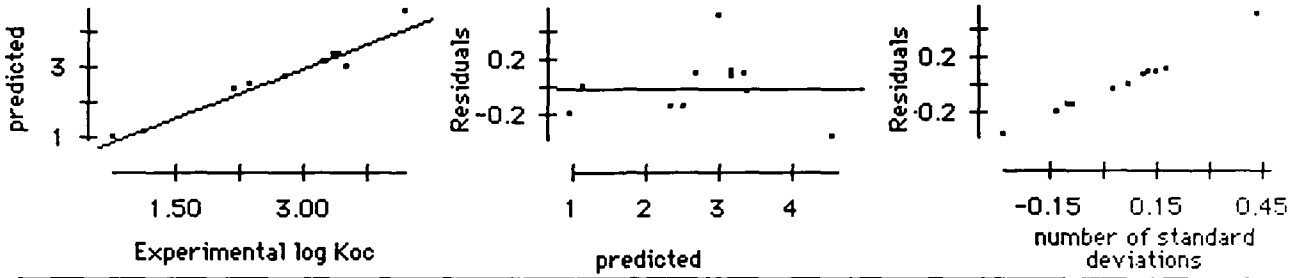
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.0615 | 0.23 | 0.267 |
| np0 | 0.2486 | 0.0234 | 10.6 |
| Δvp1 | 0.3348 | 0.0762 | 4.39 |

### Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 10.3570 | 2 | 5.178 | 80.2 |
| Residual | 0.516495 | 8 | 0.0646 | |
| Total | 10.87349 | 10 | 1.0873 | |

$r^2$ = 95.2% $n_{obs}$ = 11 s = 0.2541

Predicted vs. Exp. Residual vs. Predicted Residual vs. Prob.




Help List Return to PEP data base

## STATISTICS Class: Koc PCBs

### Regression Results

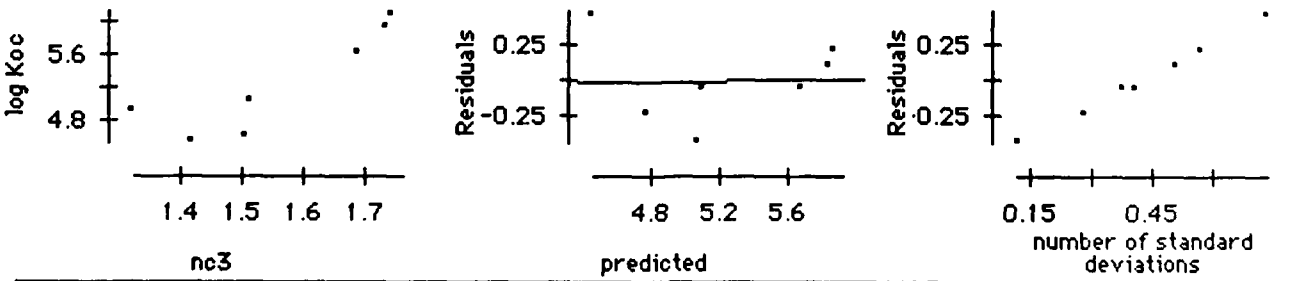
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.0018 | 1.252 | 0.001 |
| nc3 | 3.3677 | 0.799 | 4.21 |

### Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 1.84147 | 1 | 1.8415 | 17.8 |
| Residual | 0.518298 | 5 | 0.1037 | |
| Total | 2.359768 | 6 | 0.3933 | |

$r^2$ = 78.0% $n_{obs}$ = 7 s = 0.3220

Log Koc vs. MCI Residual vs. Predicted Residual vs. Prob.




Help List Return to PEP data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: Koc PAHs

### Regression Results

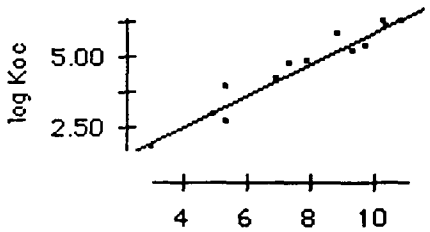
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.2187 | 0.322 | 0.679 |
| np1 | 0.5672 | 0.0405 | 14 |

### Analysis of Variance Table

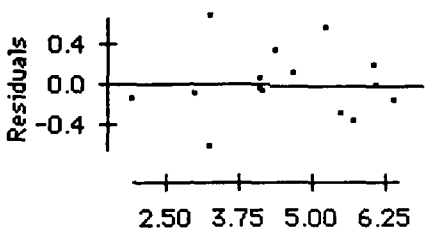
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 24.1463 | 1 | 24.15 | 196 |
| Residual | 1.60011 | 13 | 0.1231 | |
| Total | 25.74641 | 14 | 1.839 | |

$r^2$ = 93.8% $n_{obs}$ = 15 s = 0.3508

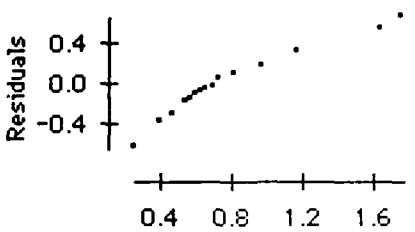
Log Koc vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

## STATISTICS Class: Koc non-polar, non-aromatic, non-

### Regression Results

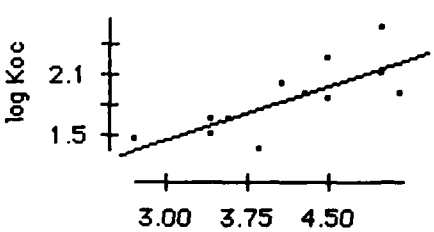
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.3614 | 0.3952 | 0.915 |
| bp0 | 0.3604 | 0.0944 | 3.82 |

### Analysis of Variance Table

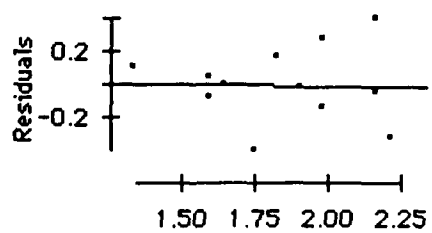
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 0.815111 | 1 | 0.8151 | 14.6 |
| Residual | 0.559114 | 10 | 0.0559 | |
| Total | 1.374225 | 11 | 0.1249 | |

$r^2$ = 59.3% $n_{obs}$ = 12 s = 0.2365

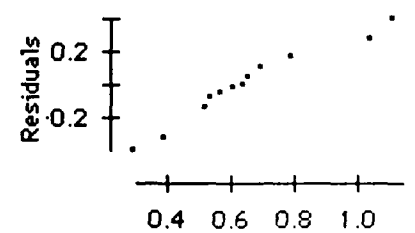
Log Koc vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: Koc non-polar, Aromatics

### Regression Results

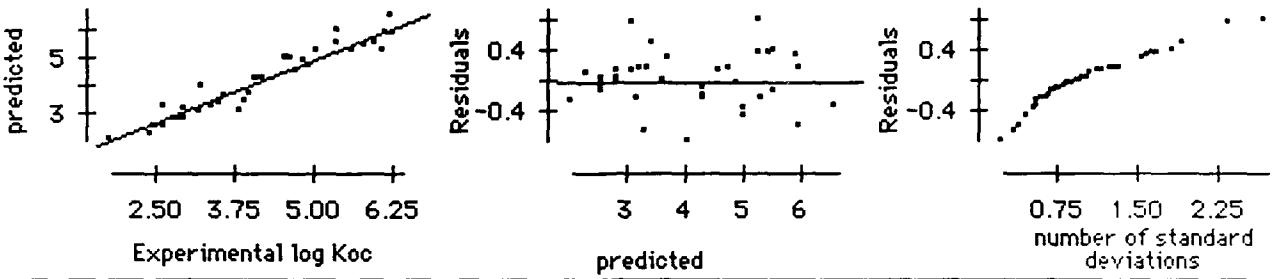
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.3889 | 0.1777 | 2.19 |
| np1 | 0.5983 | 0.0273 | 21.9 |
| bch6 | -3.4659 | 0.484 | -7.16 |

### Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 60.5243 | 2 | 30.26 | 240 |
| Residual | 4.54683 | 36 | 0.1263 | |
| Total | 65.07113 | 38 | 1.7124 | |

$r^2$ = 93.0% $n_{obs}$ = 39 s = 0.3554

Predicted vs. Exp. Residual vs. Predicted Residual vs. Prob.




Help List Return to PEP data base

## STATISTICS Class: Koc Halogenated Aromatics

### Regression Results

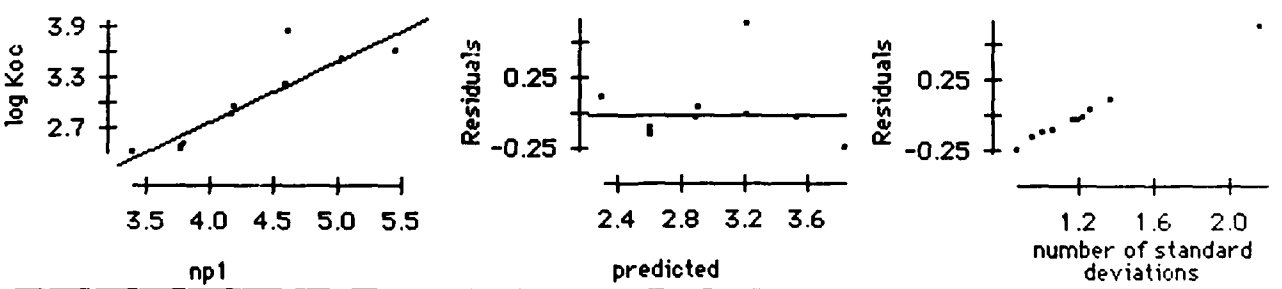
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -.2045 | 0.5708 | -.358 |
| np1 | 0.7411 | 0.1318 | 5.62 |

### Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 2.03980 | 1 | 2.0398 | 31.6 |
| Residual | 0.515841 | 8 | 0.0645 | |
| Total | 2.555641 | 9 | 0.284 | |

$r^2$ = 79.8% $n_{obs}$ = 10 s = 0.2539

Log Koc vs. MCI Residual vs. Predicted Residual vs. Prob.




Help List Return to PEP data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: Koc Carbamates

### Regression Results

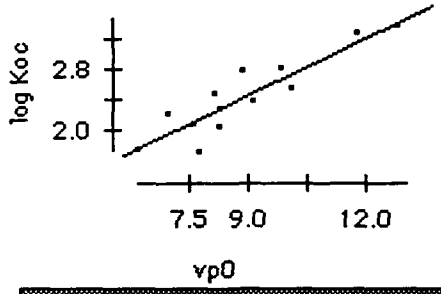
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.1814 | 0.3295 | 0.551 |
| vp0 | 0.2523 | 0.0363 | 6.96 |

### Analysis of Variance Table

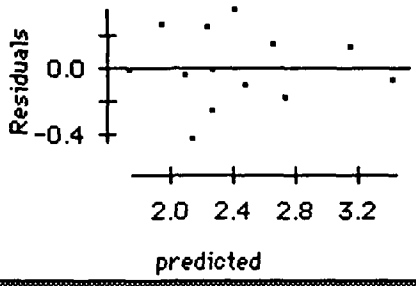
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 2.63502 | 1 | 2.635 | 48.4 |
| Residual | 0.598675 | 11 | 0.0544 | |
| Total | 3.233695 | 12 | 0.2695 | |

$r^2$ = 81.5% $n_{obs}$ = 13 s = 0.2333

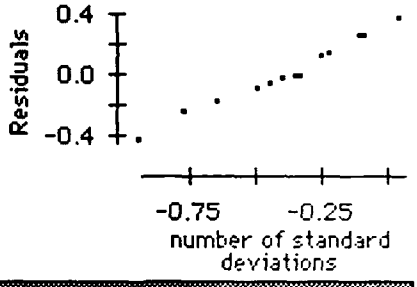
Log Koc vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

## STATISTICS Class: Koc Anilines

### Regression Results

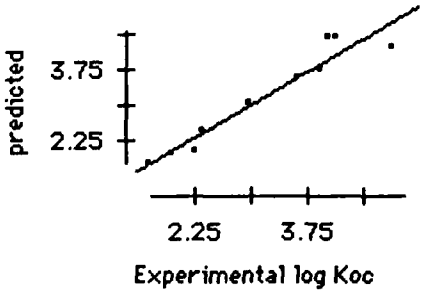
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.1424 | 0.3346 | 0.425 |
| np1 | 0.5734 | 0.0776 | 7.39 |
| Δvp0 | -1.2366 | 0.3295 | -3.75 |

### Analysis of Variance Table

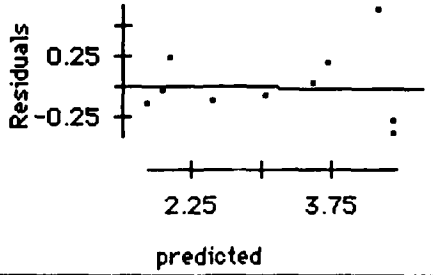
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 9.94209 | 2 | 4.971 | 45.6 |
| Residual | 0.763154 | 7 | 0.109 | |
| Total | 10.70524 | 9 | 1.1895 | |

$r^2$ = 92.9% $n_{obs}$ = 10 s = 0.3302

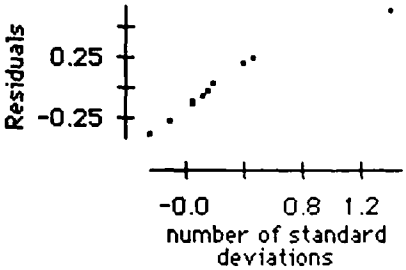
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: Koc Acetanilide

## Regression Results

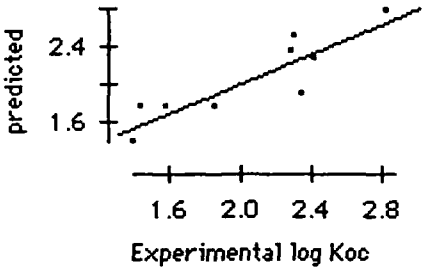
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.3745 | 0.3781 | 0.99 |
| Δvp1 | -1.2323 | 0.5152 | -2.39 |
| np1 | 0.3512 | 0.0742 | 4.74 |

## Analysis of Variance Table

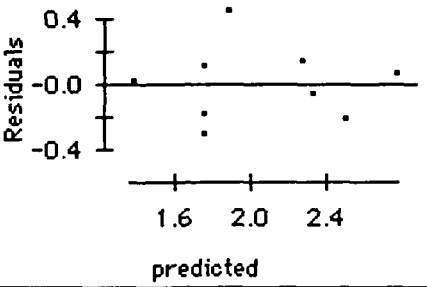
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 1.58075 | 2 | 0.7904 | 11.8 |
| Residual | 0.403537 | 6 | 0.0673 | |
| Total | 1.984287 | 8 | 0.248 | |

$r^2$ = 79.7% $n_{obs}$ = 9 s = 0.2593

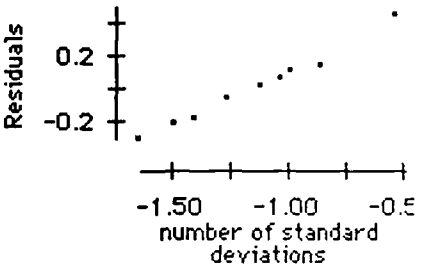
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

# STATISTICS Class: H Universal

## Regression Results

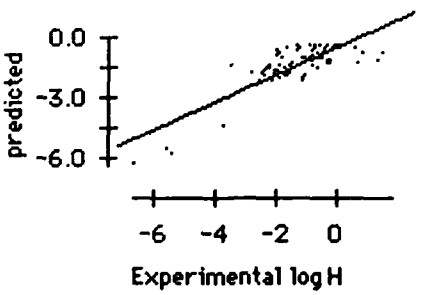
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.0003 | 0.2096 | 0.001 |
| np1 | -.2496 | 0.0422 | -5.91 |
| Δvp1 | -1.667 | 0.2082 | -8.01 |

## Analysis of Variance Table

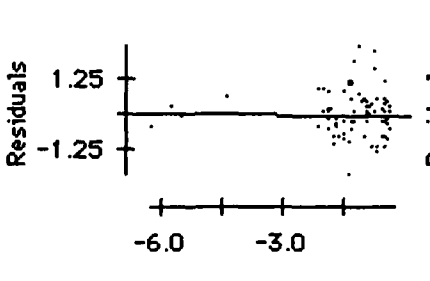
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 93.9802 | 2 | 46.99 | 74.7 |
| Residual | 45.9296 | 73 | 0.6292 | |
| Total | 139.9098 | 75 | 1.8655 | |

$r^2$ = 67.2% $n_{obs}$ = 76 s = 0.7932

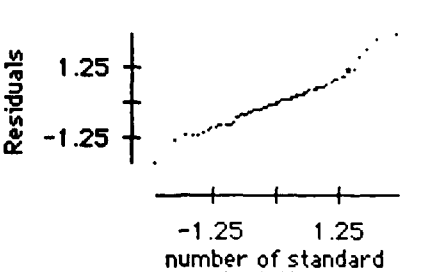
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: H nonpolar Aromatics

### Regression Results

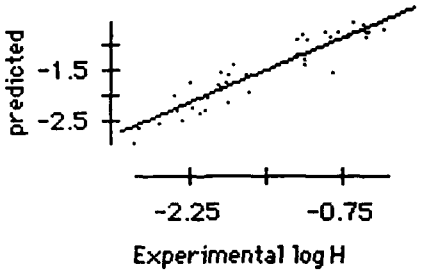
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -.9387 | 0.1182 | -7.94 |
| np6 | -1.5775 | 0.1463 | -10.8 |
| bp4 | 1.5632 | 0.209 | 7.48 |
| vp3 | -.3061 | 0.1107 | -2.76 |

### Analysis of Variance Table

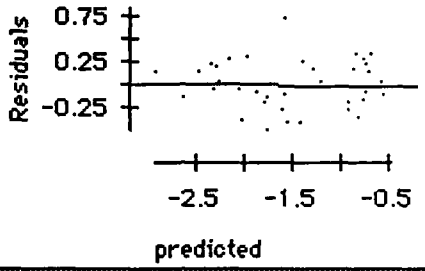
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 17.3912 | 3 | 5.797 | 79.7 |
| Residual | 2.47279 | 34 | 0.0727 | |
| Total | 19.86399 | 37 | 0.5369 | |

$r^2$= 87.6% $n_{obs}$ = 38 s = 0.2697

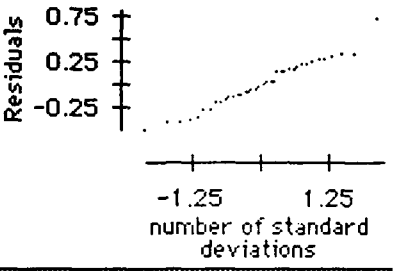
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

## STATISTICS Class: H Halogenated Aliphatics

### Regression Results

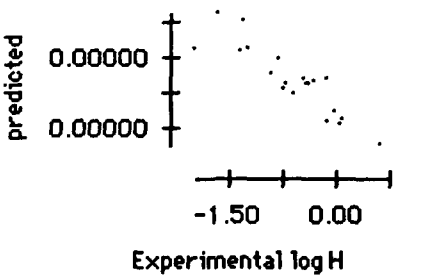
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.7325 | 0.2612 | 2.8 |
| vp1 | -1.2237 | 0.1559 | -7.85 |
| np2 | 0.6848 | 0.0961 | 7.13 |

### Analysis of Variance Table

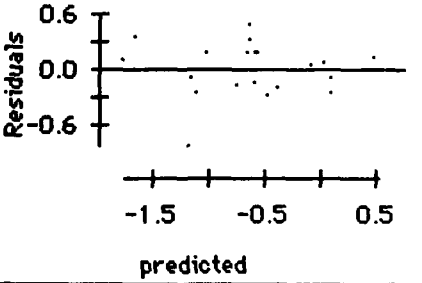
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 6.39442 | 2 | 3.1972 | 33.9 |
| Residual | 1.60216 | 17 | 0.0942 | |
| Total | 7.99658 | 19 | 0.4209 | |

$r^2$= 80.0% $n_{obs}$ = 20 s = 0.3070

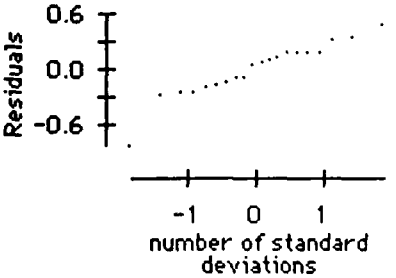
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: H PCBs

## Regression Results

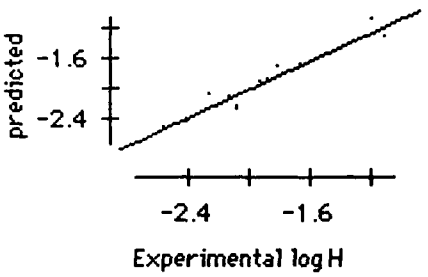
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -2.477 | 0.099 | -25 |
| npc6 | 0.3676 | 0.0247 | 14.9 |
| vpc4 | -1.1555 | 0.0895 | -12.9 |

## Analysis of Variance Table

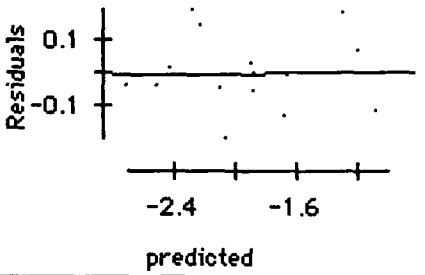
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 3.26115 | 2 | 1.6306 | 111 |
| Residual | 0.175504 | 12 | 0.0146 | |
| Total | 3.436654 | 14 | 0.2455 | |

$r^2$ = 94.9% $n_{obs}$ = 15 s = 0.1209

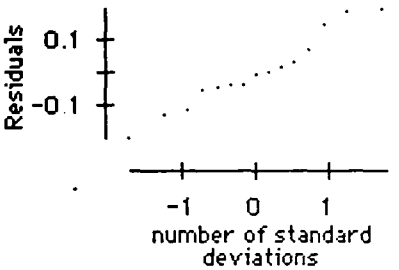
**Predicted vs. Exp.**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

# STATISTICS Class: H PAHs

## Regression Results

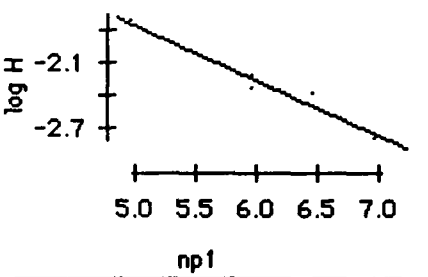
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.8541 | 0.2958 | 2.89 |
| np1 | -.5197 | 0.0495 | -10.5 |

## Analysis of Variance Table

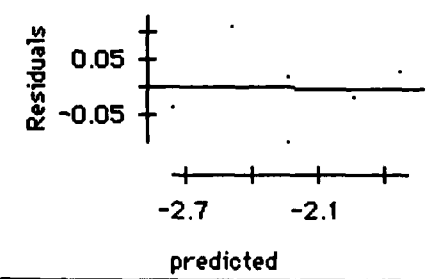
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 0.687162 | 1 | 0.6872 | 110 |
| Residual | 0.024924 | 4 | 0.0062 | |
| Total | 0.712086 | 5 | 0.1424 | |

$r^2$ = 96.5% $n_{obs}$ = 6 s = 0.0789

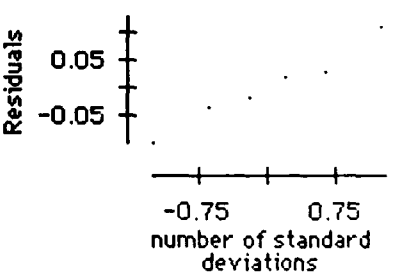
**Log H vs. MCI**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

# STATISTICS Class: BCF Universal

## Regression Results

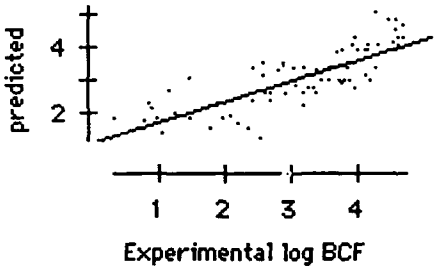
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.9816 | 0.2329 | 4.21 |
| np1 | 0.4347 | 0.0415 | 10.5 |
| Δvp0 | -1.9999 | 0.2886 | -6.93 |

## Analysis of Variance Table

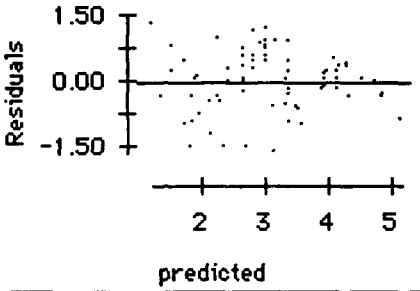
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 60.1143 | 2 | 30.06 | 61.6 |
| Residual | 32.7178 | 67 | 0.4883 | |
| Total | 92.8321 | 69 | 1.3454 | |

$r^2$ = 64.8% $n_{obs}$ = 70 s = 0.6988

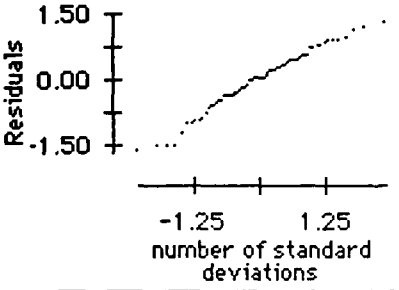
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

# STATISTICS Class: BCF PCBs

## Regression Results

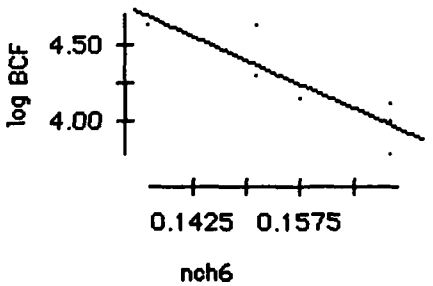
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 7.5756 | 0.6252 | 12.1 |
| nch6 | -21.162 | 4.009 | -5.28 |

## Analysis of Variance Table

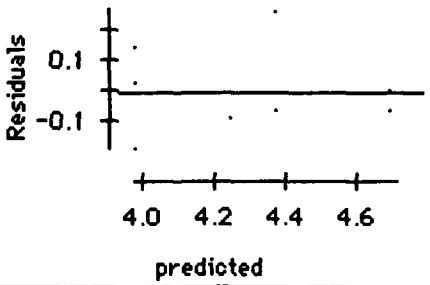
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 0.641375 | 1 | 0.6414 | 27.9 |
| Residual | 0.138126 | 6 | 0.023 | |
| Total | 0.779501 | 7 | 0.1114 | |

$r^2$ = 82.3% $n_{obs}$ = 8 s = 0.1517

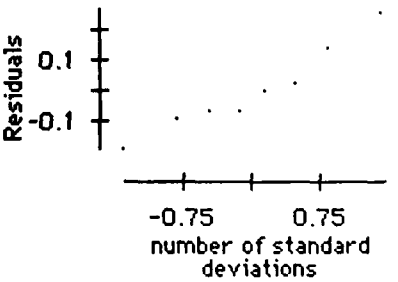
Log BCF vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

Summary Statistics Cards for MCI-property relationships

## STATISTICS Class: BCF Polar Aromatics

### Regression Results

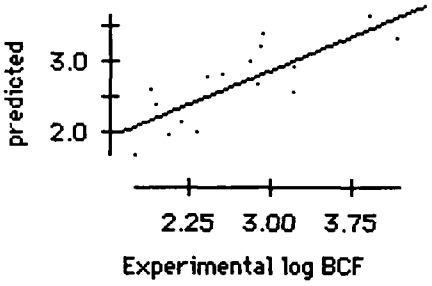
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 1.8367 | 0.5026 | 3.65 |
| Δvp0 | -1.3739 | 0.4012 | -3.42 |
| vp0 | 0.1889 | 0.0489 | 3.87 |

### Analysis of Variance Table

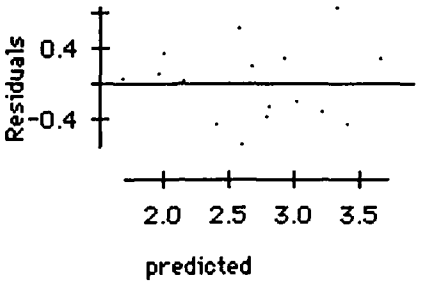
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 4.68185 | 2 | 2.3409 | 11.5 |
| Residual | 2.65010 | 13 | 0.2039 | |
| Total | 7.33195 | 15 | 0.4888 | |

$r^2$ = 63.9% $n_{obs}$ = 16 s = 0.4515

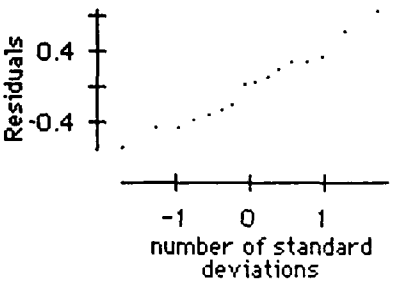
Predicted vs. Exp.




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

## STATISTICS Class: BCF Halogenated Aliphatics

### Regression Results

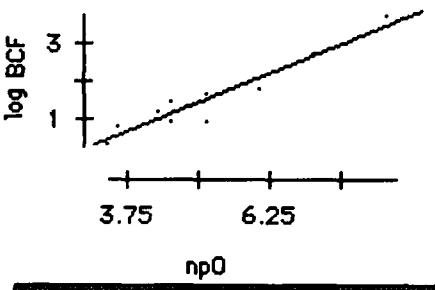
| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -1.7216 | 0.3819 | -4.51 |
| np0 | 0.632 | 0.0707 | 8.93 |

### Analysis of Variance Table

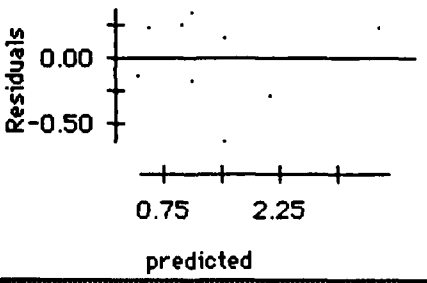
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 8.51622 | 1 | 8.5162 | 79.8 |
| Residual | 0.853741 | 8 | 0.1067 | |
| Total | 9.369961 | 9 | 1.0411 | |

$r^2$ = 90.9% $n_{obs}$ = 10 s = 0.3267

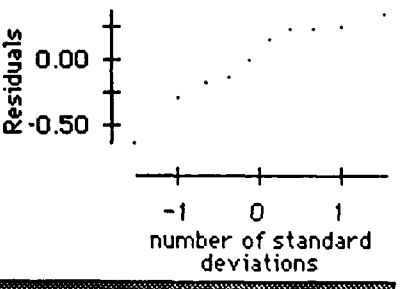
Log BCF vs. MCI




Residual vs. Predicted




Residual vs. Prob.




Help | List | Return to PEP | data base

APPENDIX B

# STATISTICS Class: S Universal from Koc

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.321341 | 0.2966 | 1.08 |
| log Koc | -1.27402 | 0.0915 | -13.9 |

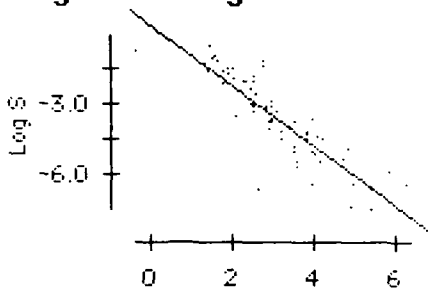
**Results in units of** mol/L

## Analysis of Variance Table

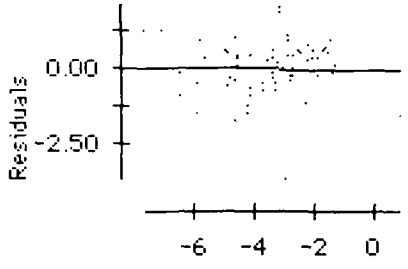
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 152.404 | 1 | 152 | 194 |
| Residual | 51.8595 | 66 | 0.78575 | |

$r^2$ = 74.6% $n_{obs}$ = 68 S = 0.8864

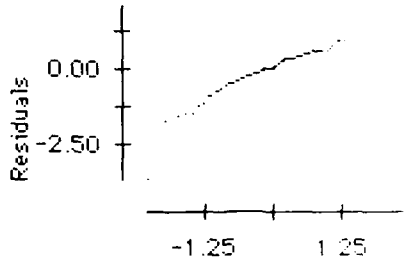
**Log S vs. Log Koc**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | Delete | New | List | data base

# STATISTICS Class: S Universal from Kow

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.73309 | 0.1511 | 4.85 |
| log Kow | -1.1348 | 0.0388 | -29.2 |

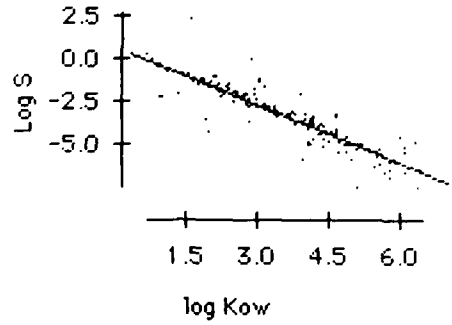
**Results in units of** mol/L

## Analysis of Variance Table

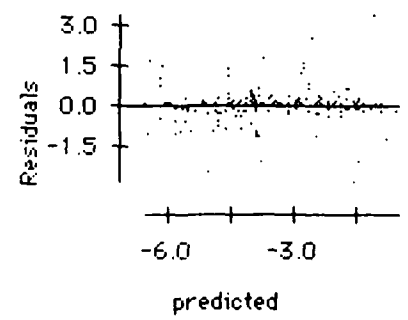
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 442.425 | 1 | 442 | 853 |
| Residual | 89.6957 | 173 | 0.518472 | |

$r^2$ = 83.1% $n_{obs}$ = 175 S = 0.7201

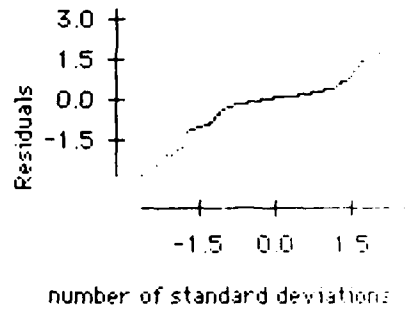
**Log S vs. Log Kow**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | Delete | New | List | data base

# STATISTICS Class: S All eq 2-15 tb 2-3 from Kow

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.978 | | |
| Log Kow | -1.339 | | |

Results in units of mol/L

## Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| | | | | |

$r^2$= 87.4 $n_{obs}$ = 156 s =

Predicted vs. Exp. Residual vs. Predicted Residual vs. Prob.

Help Delete New List data base

# STATISTICS Class: S Phosphate esters from Kow

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 12.9 | | |
| Log Kow | -2.38 | | |

Results in units of µmol/L

## Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| | | | | |

$r^2$= 65.6 $n_{obs}$ = 11 s =

Predicted vs. Exp. Residual vs. Predicted Residual vs. Prob.

Help Delete New List data base

# STATISTICS Class: S Halobenzenes from Kow & Mp

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.7178 | | |
| Log Kow | -0.9874 | | |
| Mp | -0.0095 | | |

Results in units of mol/L

## Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| | | | | |

$r^2$= 99.0 $n_{obs}$ = 35 S =

Predicted vs. Exp. Residual vs. Predicted Residual vs. Prob.

Help Delete New List data base

# STATISTICS Class: S PAHs from Kow & Mp

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| constant | -0.012 | | |
| Log Kow | -0.88 | | |
| Mp | -0.01 | | |

Results in units of mol/L

## Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| | | | | |

$r^2$= 97.9 $n_{obs}$ = 32 S =

Predicted vs. Exp. Residual vs. Predicted Residual vs. Prob.

Help Delete New List data base

# STATISTICS Class: Kow Universal from S

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 1.14855 | 0.0955 | 12.0 |
| log S (mol/L) | -0.73268 | 0.0251 | -29.2 |

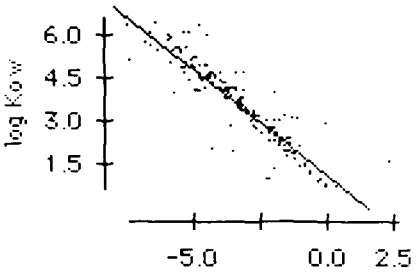
Results in units of

## Analysis of Variance Table

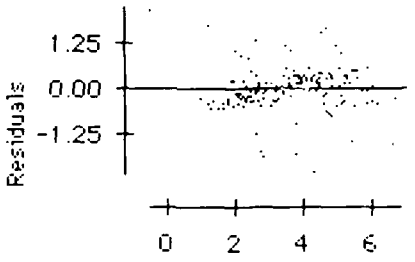
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 285.65 | 1 | 286 | 853 |
| Residual | 57.9116 | 173 | 0.33475 | |

$r^2$ = 83.1% $n_{obs}$ = 175 S = 0.5786

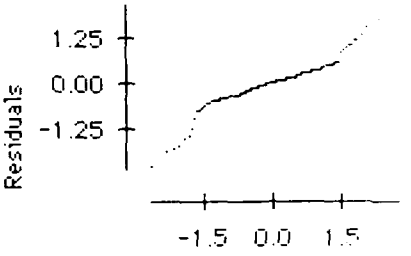
**Log Kow vs. Log S**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | Delete | New | List | data base

# STATISTICS Class: Koc Pesticides eq 4-5 from S

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 3.64 | | |
| log S mg\L | -0.55 | | |

Results in units of

## Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| | | | | |

$r^2$ = 71% $n_{obs}$ = 106 S =

**Predicted vs. Exp.**

**Residual vs. Predicted**

**Residual vs. Prob.**

Help | Delete | New | List | data base

# STATISTICS Class: Koc Universal from S

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 0.955372 | 0.1653 | 5.78 |
| log S (mol/L) | -0.58564 | 0.0421 | -13.9 |

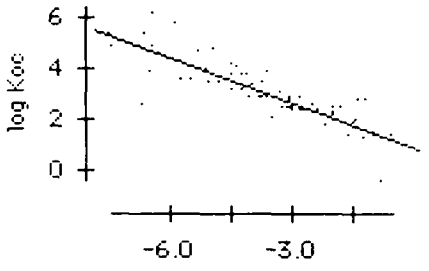
## Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 70.0563 | 1 | 70.06 | 194 |
| Residual | 23.8385 | 66 | 0.36119 | |

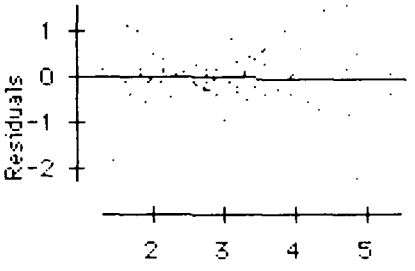
$r^2$ = 74.6%  $n_{obs}$ = 68  S = 0.601

Results in units of

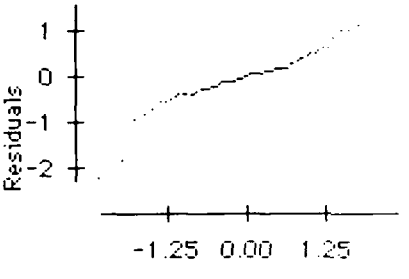
**Log Koc vs. Log S**




**Residual vs. Predicted**




**Residual vs. Prob.**




Help | Delete | New | List | data base

# STATISTICS Class: Koc Pesticides eq 4-8 tb 4-1 from Kow

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | 1.377 | | |
| log Kow | .544 | | |

## Analysis of Variance Table

| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| | | | | |

$r^2$ = .74  $n_{obs}$ = 45  S =

Results in units of

**Predicted vs. Exp.**

**Residual vs. Predicted**

**Residual vs. Prob.**

Help | Delete | New | List | data base

# STATISTICS Class: BCF Universal from Kow

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -0.58027 | 0.1978 | -2.93 |
| log Kow | 0.880366 | 0.0484 | 18.2 |

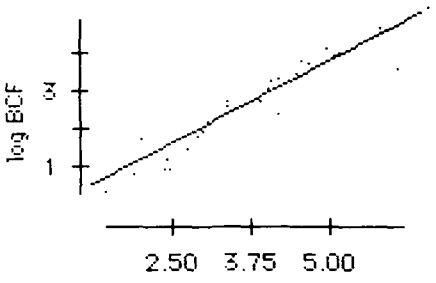
Results in units of

## Analysis of Variance Table

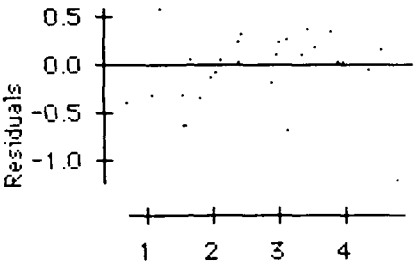
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 46.9153 | 1 | 46.92 | 331 |
| Residual | 6.51802 | 46 | 0.14170 | |

$r^2$ = 87.8% $n_{obs}$ = 48 S = 0.3764

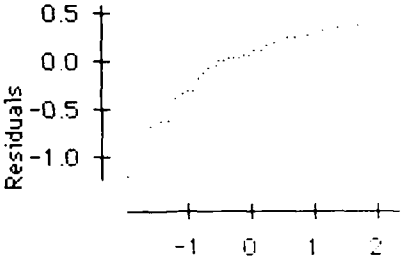
Log BCF vs. Log Kow




Residual vs. Predicted




Residual vs. Prob.




Help Delete New List data base

# STATISTICS Class: BCF Universal from Koc

## Regression Results

| Variable | Coef. | Std. Error | t |
|---|---|---|---|
| Constant | -0.81607 | 0.3946 | -2.07 |
| log Koc | 1.24095 | 0.1313 | 9.45 |

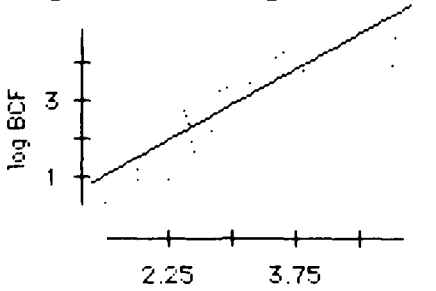
Results in units of

## Analysis of Variance Table

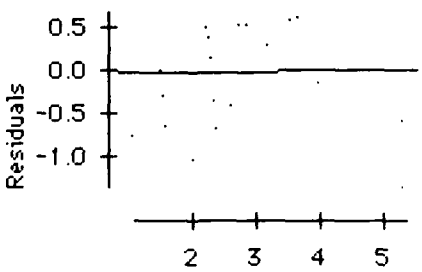
| Source | RSS | $d_f$ | MSS | F |
|---|---|---|---|---|
| Regression | 28.7887 | 1 | 28.79 | 89.3 |
| Residual | 9.02727 | 28 | 0.32240 | |

$r^2$ = 76.1% $n_{obs}$ = 30 S = 0.5678

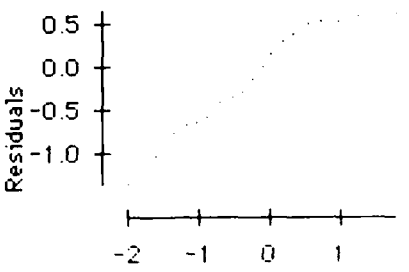
Log BCF vs. Log Koc




Residual vs. Predicted




Residual vs. Prob.




Help Delete New List data base

Subgroup name, M.W.

CH3,15.035
CH2,14.027
CH,13.019
C,12.011
CH2=CH,27.046
CH=CH,26.038
CH2=C,26.038
CH=C,25.03
C=C,24.022
ACH,13.019
AC,12.011
ACCH3,27.046
ACCH2,26.038
ACCH,25.03
OH,17.0073
CH3OH,32.0423
H2O,18.0152
ACOH,29.0183
CH3CO,43.0454
CH2CO,42.0375
CHO,25.03
CH3COO,59.0448
CH2COO,58.0369
HCOO,41.0294
CH3O,31.0344
CH2O,30.0265
CH-O,29.0186
FCH2O,49.0249
CH3NH2,31.0575
CH2NH2,30.0495
CHNH2,29.0415
CH3NH,30.0496
CH2NH,29.0416
CHNH,28.0336
CH3N,29.0417
CH2N,28.0337
ACNH2,28.0337
C5H5N,79.1012
C5H4N,78.0933
C5H3N,77.0854
CH3CN,41.0527
CH2CN,40.0447
COOH,45.0167
HCOOH,46.0246
CH2Cl,49.48
CHCl,48.472
CCl,47.464
CH2Cl2,84.933
CHCl2,83.925
CCl2,82.917
CHCl3,119.378
CCl3,118.37
CCl4,153.823
ACCl,47.464
CH3NO2,61.0405
CH2NO2,60.0325
CHNO2,59.0245
ACNO2,58.0165
CS2,76.131
CH3SH,48.1029
CH2SH,47.0949
Furfural,96.08
(CH2OH)2,62.0686
I,126.9045
Br,79.904
CH≡C,25.0299
C≡C,24.022
DMSO,78.13
ACRY,53.06
Cl-(C=C),59.475
ACF,31.0094
DMF-1,
DMF-2,
CF3,69.0062
CF2,50.0078
CF,31.0094

Volume & Area

Import

Export